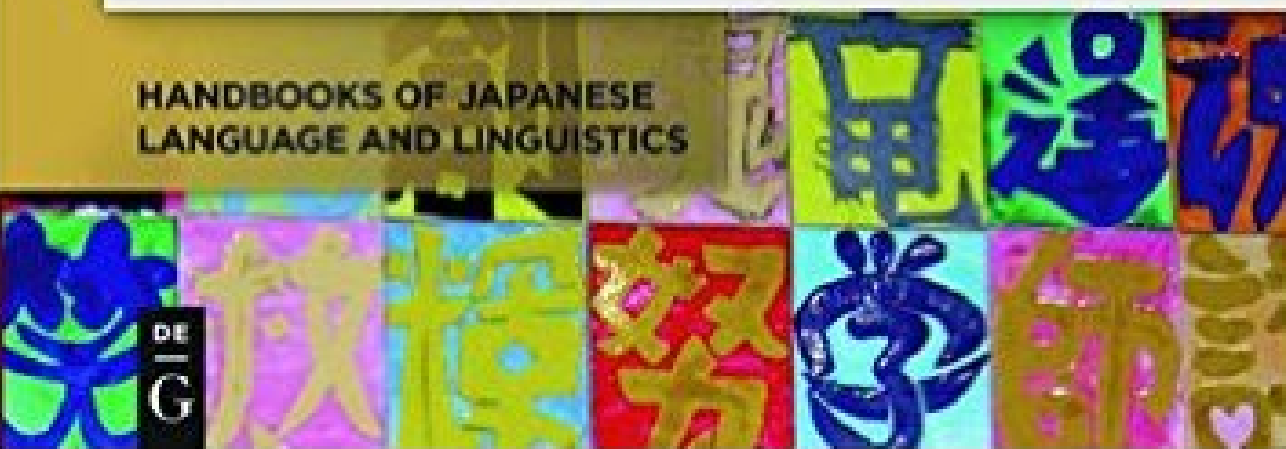*Taro Kageyama,*
Hideki Kishimoto (Eds.)

HANDBOOK OF JAPANESE LEXICON AND WORD FORMATION

NINJAL
National Institute for Japanese Language and Linguistics

HANDBOOKS OF JAPANESE LANGUAGE AND LINGUISTICS

Taro Kageyama and Hideki Kishimoto (Eds.)
Handbook of Japanese Lexicon and Word Formation

Handbooks of Japanese Language and Linguistics

Edited by
Masayoshi Shibatani
Taro Kageyama

Volume 3

Handbook of Japanese Lexicon and Word Formation

Edited by
Taro Kageyama
Hideki Kishimoto

DE GRUYTER
MOUTON

Preface

The project of compiling a series of comprehensive handbooks covering major fields of Japanese linguistics started in 2011, when Masayoshi Shibatani received a commission to edit such volumes as series editor from De Gruyter Mouton. As the planning progressed, with the volume titles selected and the volume editors assigned, the enormity of the task demanded the addition of a series co-editor. Taro Kageyama, Director-General of the National Institute for Japanese Language and Linguistics (NINJAL), was invited to join the project as a series co-editor. His participation in the project opened the way to make it a joint venture between NINJAL and De Gruyter Mouton. We are pleased to present the *Handbooks of Japanese Language and Linguistics* (*HJLL*) as the first materialization of the agreement of academic cooperation concluded between NINJAL and De Gruyter Mouton.

The HJLL Series is composed of twelve volumes, primarily focusing on Japanese but including volumes on the Ryukyuan and Ainu languages, which are also spoken in Japan, as well as some chapters on Japanese Sign Language in the applied linguistics volume.

– Volume 1: *Handbook of Japanese Historical Linguistics*
– Volume 2: *Handbook of Japanese Phonetics and Phonology*
– Volume 3: *Handbook of Japanese Lexicon and Word Formation*
– Volume 4: *Handbook of Japanese Syntax*
– Volume 5: *Handbook of Japanese Semantics and Pragmatics*
– Volume 6: *Handbook of Japanese Contrastive Linguistics*
– Volume 7: *Handbook of Japanese Dialects*
– Volume 8: *Handbook of Japanese Sociolinguistics*
– Volume 9: *Handbook of Japanese Psycholinguistics*
– Volume 10: *Handbook of Japanese Applied Linguistics*
– Volume 11: *Handbook of the Ryukyuan Languages*
– Volume 12: *Handbook of the Ainu Language*

Surpassing all currently available reference works on Japanese in both scope and depth, the *HJLL* series provides a comprehensive survey of nearly the entire field of Japanese linguistics. Each volume includes a balanced selection of articles contributed by established linguists from Japan as well as from outside Japan and is critically edited by volume editors who are leading researchers in their individual fields. Each article reviews milestone achievements in the field, provides an overview of the state of the art, and points to future directions of research. The twelve titles are thus expected individually and collectively to contribute not only to the enhancement of studies on Japanese on the global level but also to the opening up of new perspectives for general linguistic research from both empirical and theoretical standpoints.

The *HJLL* project has been made possible by the active and substantial participation of numerous people including the volume editors and authors of individual

chapters. We would like to acknowledge with gratitude the generous support, both financial and logistic, given to this project by NINJAL. We are also grateful to John Haig (retired professor of Japanese linguistics, the University of Hawai'i at Mānoa), serving as copy-editor for the series. In the future, more publications are expected to ensue from the NINJAL-Mouton academic cooperation.

Masayoshi Shibatani, Deedee McMurtry Professor of Humanities and Professor of Linguistics, Rice University/Professor Emeritus, Kobe University
Taro Kageyama, Director-General, National Institute for Japanese Language and Linguistics (NINJAL)/Professor Emeritus, Kwansei Gakuin University

Masayoshi Shibatani and Taro Kageyama

Introduction to the *Handbooks of Japanese Language and Linguistics*

Comprising twelve substantial volumes, the *Handbooks of Japanese Language and Linguistics* (HJLL) series provides a comprehensive survey of practically all the major research areas of Japanese linguistics on an unprecedented scale, together with surveys of the endangered languages spoken in Japan, Ryukyuan and Ainu. What follows are introductions to the individual handbooks, to the general conventions adopted in this series, and the minimum essentials of contemporary Standard Japanese. Fuller descriptions of the languages of Japan, Japanese grammar, and the history of the Japanese language are available in such general references as Martin (1975), Shibatani (1990), and Frellesvig (2010).

1 Geography, Population, and Languages of Japan

Japan is situated in the most populous region of the world – Asia, where roughly one half of the world population of seven billion speak a variety of languages, many of which occupy the top tier of the ranking of the native-speaker population numbers. Japanese is spoken by more than 128 million people (as of 2013), who live mostly in Japan but also in Japanese emigrant communities around the world, most notably Hawaii, Brazil and Peru. In terms of the number of native speakers, Japanese ranks ninth among the world's languages. Due partly to its rich and long literary history, Japanese is one of the most intensely studied languages in the world and has received scrutiny under both the domestic grammatical tradition and those developed outside Japan such as the Chinese philological tradition, European structural linguistics, and generative grammar developed in America. The *Handbooks of Japanese Language and Linguistics* intend to capture the achievements garnered over the years through analyses of a wide variety of phenomena in a variety of theoretical frameworks.

As seen in Map 1, where Japan is shown graphically superimposed on Continental Europe, the Japanese archipelago has a vast latitudinal extension of approximately 3,000 kilometers ranging from the northernmost island, roughly corresponding to Stockholm, Sweden, to the southernmost island, roughly corresponding to Sevilla, Spain.

Contrary to popular assumption, Japanese is not the only language native to Japan. The northernmost and southernmost areas of the Japanese archipelago are inhabited by people whose native languages are arguably distinct from Japanese. The southernmost sea area in Okinawa Prefecture is dotted with numerous small islands
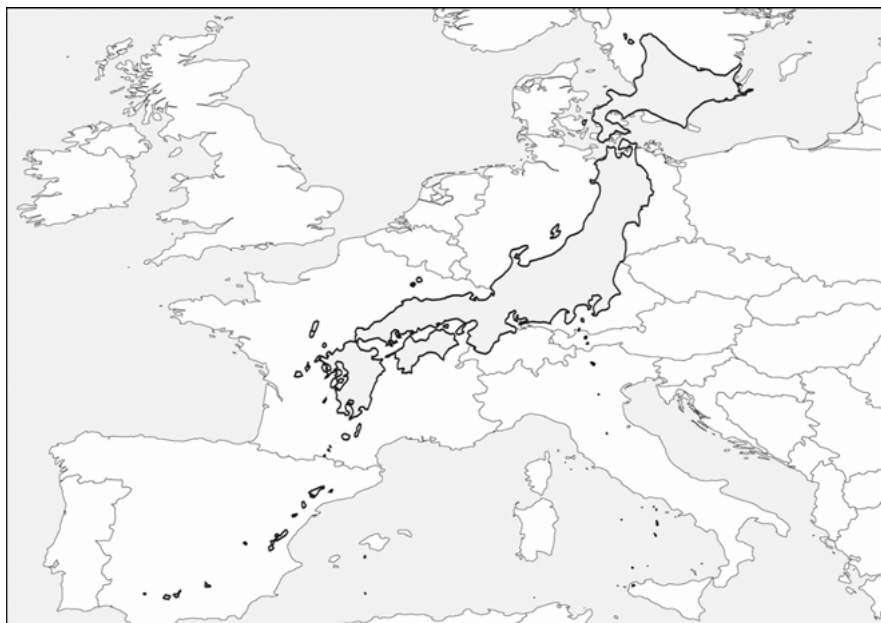
Map 1: *Japan as overlaid on Europe*
Source: Shinji Sanada. 2007. *Hōgen wa kimochi o tsutaeru [Dialects convey your heart]*.
Tokyo: Iwanami, p. 68.

where Ryukyuan languages are spoken. Until recent years, Japanese scholars tended to treat Ryukyuan language groups as dialects of Japanese based on fairly transparent correspondences in sounds and grammatical categories between mainland Japanese and Ryukyuan, although the two languages are mutually unintelligible. Another reason that Ryukyuan languages have been treated as Japanese dialects is that Ryukyuan islands and Japan form a single nation. In terms of nationhood, however, Ryukyu was an independent kingdom until the beginning of the seventeenth century, when it was forcibly annexed to the feudal domain of Satsuma in southern Kyushu.

A more recent trend is to treat Ryukyuan as forming a branch of its own with the status of a sister language to Japanese, following the earlier proposals by Chamberlain (1895) and Miller (1971). Many scholars specializing in Ryukyuan today even confer language status to different language groups within Ryukyuan, such as Amami language, Okinawan language, Miyako language, etc., which are grammatically distinct to the extent of making them mutually unintelligible. The prevailing view now has Japanese and Ryukyuan forming the Japonic family as daughter languages of Proto-Japonic. HJLL follows this recent trend of recognizing Ryukyuan as a sister language to Japanese and devotes one full volume to it. The ***Handbook of the Ryukyuan Languages*** provides the most up-to-date answers pertaining to Ryukyuan

language structures and use, and the ways in which these languages relate to Ryukyuan society and history. Like all the other handbooks in the series, each chapter delineates the boundaries and the research history of the field it addresses, comprises the most important and representative information on the state of research, and spells out future research desiderata. This volume also includes a comprehensive bibliography of Ryukyuan linguistics.

The situation with Ainu, another language indigenous to Japan, is much less clear as far as its genealogy goes. Various suggestions have been made relating Ainu to Paleo-Asiatic, Ural-Altaic, and Malayo-Polynesian or to such individual languages as Gilyak and Eskimo, besides the obvious candidate of Japanese as its sister language. The general consensus, however, points to the view that Ainu is related to Japanese quite indirectly, if at all, via the Altaic family with its Japanese-Korean subbranch (see Miller 1971; Shibatani 1990: 5–7 for an overview). Because Ainu has had northern Japan as its homeland and because HJLL is also concerned with various aspects of Japanese linguistics scholarship in general, we have decided to include a volume devoted to Ainu in this series. The **Handbook of the Ainu Language** outlines the history and current state of the Ainu language, offers a comprehensive survey of Ainu linguistics, describes major Ainu dialects in Hokkaido and Sakhalin, and devotes a full section to studies dealing with typological characteristics of the Ainu language such as polysynthesis and incorporation, person marking, plural verb forms, and aspect and evidentials.

2 History

Japan's rich and long literary history dates back to the seventh century, when the Japanese learned to use Chinese characters in writing Japanese. Because of the availability of abundant philological materials, the history of the Japanese language has been one of the most intensely pursued fields in Japanese linguistics. While several different divisions of Japanese language history have been proposed, Frellesvig (2010) proposes the following four linguistic periods, each embracing the main political epochs in Japanese history.

1. Old Japanese   700–800   (Nara period, 712–794)
2. Early Middle Japanese 800–1200   (Heian period, 794–1185)
3. Late Middle Japanese   1200–1600   (Kamakura period, 1185–1333;
                                        Muromachi period, 1333–1573)
4. Modern Japanese   1600–   (Edo, 1603–1868; Meiji, 1868–1912;
                                        Taishō, 1912–1926; Shōwa, 1926–1989;
                                        Heisei, 1989–)

This division reflects a major gulf between Pre-modern and Modern Japanese caused by some radical changes in linguistic structure during the Late Middle Japanese period. Modern Japanese is often further subdivided into Early Modern (Edo, 1603–1868), Modern (Meiji, 1868–1912; Taishō, 1912–1926), and Present-day Japanese (Shōwa, 1926–1989; Heisei, 1989–).

The ***Handbook of Japanese Historical Linguistics*** will present the latest research on better studied topics, such as segmental phonology, accent, morphology, and some salient syntactic phenomena such as focus constructions. It will also introduce areas of study that have traditionally been underrepresented, ranging from syntax and Sinico-Japanese (*kanbun*) materials to historical pragmatics, and demonstrate how they contribute to a fuller understanding of the overall history of Japanese, as well as outlining larger-scale tendencies and directions in changes that have taken place within the language over its attested history. Major issues in the reconstruction of prehistoric Japanese and in the individual historical periods from Old Japanese to Modern Japanese are discussed including writing and the materials for historical studies, influences of Sinico-Japanese on Japanese, the histories of different vocabulary strata, the history of honorifics and polite language, generative diachronic syntax, and the development of case marking.

3 Geographic and Social Variations

Because of the wide geographical spread of the Japanese archipelago from north to south, characterized by high mountain ranges, deep valleys, and wide rivers as well as numerous islands, Japanese has developed a multitude of dialects, many of which differ from each other in a way more or less like current descendants of the Romance language family. Like the historical studies, the research tradition of dialect studies has a unique place in Japanese linguistics, which has also attracted a large number of students, amateur collectors of dialect forms as well as professional linguists. The ***Handbook of Japanese Dialects*** surveys the historical backdrop of the theoretical frameworks of contemporary studies in Japanese geolinguistics and includes analyses of prominent research topics in cross-dialectal perspectives, such as accentual systems, honorifics, verbs of giving, and nominalizations. The volume also devotes large space to sketch grammars of dialects from the northern island of Hokkaido to the southern island of Kyushu, allowing a panoramic view of the differences and similarities in the representative dialects throughout Japan.

Besides the physical setting fostering geographic variations, Japanese society has experienced several types of social structure over the years, starting from the time of the nobility and court life of the Old and Early Middle Japanese periods, through the caste structure of the feudalistic Late Middle and Early Modern Japanese periods, to the modern democratic society in the Modern and Present-day Japanese

periods. These different social structures spawned a variety of social dialects including power- and gender-based varieties of Japanese. The **Handbook of Japanese Sociolinguistics** examines a wide array of sociolinguistic topics ranging from the history of Japanese sociolinguistics, including foreign influences and internal innovations, to the central topics of variations due to social stratification, gender differences, and discourse genre. Specific topics include honorifics and women's speech, critical discourse analysis, pragmatics of political discourse, contact-induced change, emerging new dialects, Japanese language varieties outside Japan, and language policy.

4 Lexicon and Phonology

The literary history of Japan began with early contacts with China. Chinese apparently began to enrich the Japanese lexicon in even pre-historic periods, when such deeply assimilated words as *uma* 'horse' and *ume* 'plum' are believed to have entered the language. Starting in the middle of the sixth century, when Buddhism reached Japan, Chinese, at different periods and from different dialect regions, has continuously contributed to Japanese in an immeasurable way affecting all aspects of grammar, but most notably the lexicon and the phonological structure, which have sustained further and continuous influences from European languages from the late Edo period on. Through these foreign contacts, Japanese has developed a complex vocabulary system that is composed of four lexical strata, each with unique lexical, phonological, and grammatical properties: native Japanese, mimetic, Sino-Japanese, and foreign (especially English).

The **Handbook of Japanese Lexicon and Word Formation** presents a comprehensive survey of the Japanese lexicon, word formation processes, and other lexical matters seen in the four lexical strata of contemporary Japanese. The agglutinative character of the language, coupled with the intricate system of vocabulary strata, makes it possible for compounding, derivation, conversion, and inflection to be closely intertwined with syntactic structure, giving rise to theoretically intriguing interactions of word formation processes and syntax that are not easily found in inflectional, isolating, or polysynthetic types of languages. The theoretically oriented studies associated with these topics are complemented by those oriented toward lexical semantics, which also bring to light theoretically challenging issues involving the morphology-syntax interface.

The four lexical strata characterizing the Japanese lexicon are also relevant to Japanese phonology as each stratum has some characteristic sounds and sound combinations not seen in the other strata. The **Handbook of Japanese Phonetics and Phonology** describes and analyzes the basic phonetic and phonological structures of modern Japanese with main focus on standard Tokyo Japanese, relegating the topics of dialect phonetics and phonology to the *Handbook of Japanese Dialects*.

The handbook includes several chapters dealing with phonological processes unique to the Sino-Japanese and foreign strata as well as to the mimetic stratum. Other topics include word tone/accent, mora-timing, sequential voicing (*rendaku*), consonant geminates, vowel devoicing and diphthongs, and the appearance of new consonant phonemes. Also discussed are phonetic and phonological processes within and beyond the word such as rhythm, intonation, and the syntax-phonology interface, as well as issues bearing on other subfields of linguistics such as historical and corpus linguistics, L1 phonology, and L2 research.

5 Syntax and Semantics

Chinese loans have also affected Japanese syntax, though the extent is unclear to which they affected Japanese semantics beyond the level of lexical semantics. In particular, Chinese loans form two distinct lexical categories in Japanese – verbal nouns, forming a subcategory of the noun class, and adjectival nouns (*keiyō dōshi*), which are treated as forming major lexical categories, along with noun, verb, and adjective classes, by those who recognize this as an independent category. The former denote verbal actions, and, unlike regular nouns denoting objects and thing-like entities, they can function as verbs by combining with the light verb *suru*, which is obviously related to the verb *suru* 'do'. The nominal-verbal Janus character of verbal nouns results in two widely observed syntactic patterns that are virtually synonymous in meaning; e.g., *benkyoo-suru* (studying-DO) 'to study' and *benkyoo o suru* (studying ACC do) 'do studying'. As described in the *Handbook of Japanese Lexicon and Word Formation*, the lexical category of adjectival noun has been a perennial problem in the analysis of Japanese parts of speech. The property-concept words, e.g., *kirei* 'pretty', *kenkoo* 'health/healthy', falling in this class do not inflect by themselves unlike native Japanese adjectives and, like nouns, require the inflecting copula *da* in the predication function – hence the label of adjectival noun for this class. However, many of them cannot head noun phrases – the hallmark of the nominal class – and some of them even yield nouns via *-sa* nominalization, which is not possible with regular nouns.

The Lexicon-Word Formation handbook and the ***Handbook of Japanese Syntax*** make up twin volumes because many chapters in the former deal with syntactic phenomena, as the brief discussion above on the two Sino-Japanese lexical categories clearly indicates. The syntax handbook covers a vast landscape of Japanese syntax from three theoretical perspectives: (1) traditional Japanese grammar, known as *kokugogaku* (lit. national-language study), (2) the functional approach, and (3) the generative grammar framework. Broad issues analyzed include sentence types and their interactions with grammatical verbal categories, grammatical relations (topic, subject, etc.), transitivity, nominalization, grammaticalization, voice (passives and

causatives), word order (subject, scrambling, numeral quantifier, configurationality), case marking (*ga/no* conversion, morphology and syntax), modification (adjectives, relative clause), and structure and interpretation (modality, negation, prosody, ellipsis). These topics have been pursued vigorously over many years under different theoretical persuasions and have had important roles in the development of general linguistic theory. For example, the long sustained studies on the grammatical of subject and topic in Japanese have had significant impacts on the study of grammatical relations in European as well as Austronesian languages. In the study of word order, the analysis of Japanese numeral quantifiers is used as one of the leading pieces of evidence for the existence of a movement rule in human language. Under case marking, the way subjects are case-marked in Japanese has played a central role in the study of case marking in the Altaic language family. Recent studies of nominalizations have been central to the analysis of their modification and referential functions in a wide variety of languages from around the globe with far-reaching implications to past studies of such phenomena as parts of speech, (numeral) classifiers, and relative clauses. And the study of how in Japanese prosody plays a crucial role in interpretation has become the basis of some important recent developments in the study of wh-questions.

The **Handbook of Japanese Semantics and Pragmatics** presents a collection of studies on linguistic meaning in Japanese, either as conventionally encoded in linguistic form (the field of semantics) or as generated by the interaction of form with context (the field of pragmatics). The studies are organized around a model that has long currency in traditional Japanese grammar, whereby the linguistic clause consists of a multiply nested structure centered in a propositional core of objective meaning around which forms are deployed that express progressively more subjective meaning as one moves away from the core toward the periphery of the clause. Following this model, the topics treated in this volume range from aspects of meaning associated with the propositional core, including elements of meaning structured in lexical units (lexical semantics), all the way to aspects of meaning that are highly subjective, being most grounded in the context of the speaker. In between these two poles of the semantics-pragmatics continuum are elements of meaning that are defined at the level of propositions as a whole or between different propositions (propositional logic) and forms that situate propositions in time as events and those situating events in reality including non-actual worlds, e.g., those hoped for (desiderative meaning), denied (negation), hypothesized (conditional meaning), or viewed as ethically or epistemologically possible or necessary (epistemic and deontic modality). Located yet closer to the periphery of the Japanese clause are a rich array of devices for marking propositions according to the degree to which the speaker is committed to their veracity, including means that mark differing perceptual and cognitive modalities and those for distinguishing information variously presupposed.

These studies in Japanese syntax and semantics are augmented by cross-linguistic studies that examine various topics in these fields from the perspectives of language

universals and the comparative study of Japanese and another language. The ***Handbook of Japanese Contrastive Linguistics*** sets as its primary goal uncovering principled similarities and differences between Japanese and other languages around the globe and thereby shedding new light on the universal and language-particular properties of Japanese. Topics ranging from inalienable possession to numeral classifiers, from spatial deixis to motion typology, and from nominalization to subordination, as well as topics closely related to these phenomena are studied in the typological universals framework. Then various aspects of Japanese such as resultative-progressive polysemy, entailment of event realization, internal-state predicates, topic constructions, and interrogative pronouns, are compared and contrasted with individual languages including Ainu, Koryak, Chinese, Korean, Newar, Thai, Burmese, Tagalog, Kapampangan, Lamaholot, Romanian, French, Spanish, German, English, Swahili, Sidaama, and Mayan languages.

6 Psycholinguistics and Applied Linguistics

HJLL includes two volumes containing topics related to wider application of Japanese linguistics and to those endeavors seeking grammar-external evidence for the psycho-neurological reality of the structure and organization of grammar. By incorporating the recent progress in the study of the cognitive processes and brain mechanisms underlying language use, language acquisition, and language disorder, the ***Handbook of Japanese Psycholinguistics*** discusses the mechanisms of language acquisition and language processing. In particular, the volume seeks answers to the question of how Japanese is learned/acquired as a first or second language, and pursues the question of how we comprehend and produce Japanese sentences. The chapters in the acquisition section allow readers to acquaint themselves with issues pertaining to the question of how grammatical features (including pragmatic and discourse features) are acquired and how our brain develops in the language domain, with respect to both language-particular and universal features. Specific topics dealt with include Japanese children's perceptual development, the conceptual and grammatical development of nouns, Japanese specific language impairment, narrative development in the L1 cognitive system, L2 Japanese acquisition and its relation to L1 acquisition. The language processing section focuses on both L1 and L2 Japanese processing and covers topics such as the role of prosodic information in production/comprehension, the processing of complex grammatical structures such as relative clauses, the processing issues related to variable word order, and lexical and sentence processing in L2 by speakers of a different native language.

The ***Handbook of Japanese Applied Linguistics*** complements the Psycholinguistics volume by examining language acquisition from broader sociocultural per-

spectives, i.e., language as a means of communication and social behavioral system, emphasizing pragmatic development as central to both L1 and L2 acquisition and overall language/human development. Topics approached from these perspectives include the role of caregiver's speech in early language development, literacy acquisition, and acquisition of writing skills. Closely related to L1 and L2 acquisition/ development are studies of bilingualism/multilingualism and the teaching and learning of foreign languages, including Japanese as a second language, where topics discussed include cross-lingual transfer from L1 to L2, learning errors, and proficiency assessment of second language acquisition. Chapters dealing with topics more squarely falling in the domain of applied linguistics cover the issues in corpus/ computational linguistics (including discussions on CHILDES for Japanese and the KY corpus widely-used in research on Japanese as a second language), clinical linguistics (including discussions on language development in children with hearing impairment and other language disorders, with Down syndrome, or autism), and translation and interpretation. Technically speaking, Japanese Sign Language is not a variety of Japanese. However, in view of the importance of this language in Japanese society and because of the rapid progress in sign language research in Japan and abroad and what it has to offer to the general theory of language, chapters dealing with Japanese Sign Language are also included in this volume.

7 Grammatical Sketch of Standard Japanese

The following pages offer a brief overview of Japanese grammar as an aid for a quick grasp of the structure of Japanese that may prove useful in studying individual, thematically organized handbooks of this series. One of the difficult problems in presenting non-European language materials using familiar technical terms derived from the European grammatical tradition concerns mismatches between what the glosses may imply and what grammatical categories they are used to denote in the description. We will try to illustrate this problem below as a way of warning not to take all the glosses at their face value. But first some remarks are in order about the conventions of transcription of Japanese, glossing of examples, and their translations used in this series.

7.1 Writing, alphabetic transcription, and pronunciation

Customarily, Japanese is written by using a mixture of Chinese characters (for content words), *hiragana* (for function words such as particles, suffixes and inflectional endings), *katakana* (for foreign loans and mimetics), and sometimes Roman alphabet.

Because Japanese had no indigenous writing system, it developed two phonogram systems of representing a phonological unit of "mora", namely *hiragana* and *katakana*, by simplifying or abbreviating (parts of) Chinese characters. *Hiragana* and *katakana* syllabaries are shown in Table 1, together with the alphabetic transcriptions adopted in the HJLL series.

Table 1: *Alphabetic transcriptions adopted in HJLL*

transcription	a	ka	sa	ta	na	ha	ma	ya	ra	wa	n
hiragana	あ	か	さ	た	な	は	ま	や	ら	わ	ん
katakana	ア	カ	サ	タ	ナ	ハ	マ	ヤ	ラ	ワ	ン
transcription	i	ki	si	ti	ni	hi	mi	–	ri	–	
hiragana	い	き	し	ち	に	ひ	み	–	り	–	
katakana	イ	キ	シ	チ	ニ	ヒ	ミ	–	リ	–	
transcription	u	ku	su	tu	nu	hu	mu	yu	ru	–	
hiragana	う	く	す	つ	ぬ	ふ	む	ゆ	る	–	
katakana	ウ	ク	ス	ツ	ヌ	フ	ム	ユ	ル	–	
transcription	e	ke	se	te	ne	he	me	–	re	–	
hiragana	え	け	せ	て	ね	へ	め	–	れ	–	
katakana	エ	ケ	セ	テ	ネ	ヘ	メ	–	レ	–	
transcription	o	ko	so	to	no	ho	mo	yo	ro	o	
hiragana	お	こ	そ	と	の	ほ	も	よ	ろ	を	
katakana	オ	コ	ソ	ト	ノ	ホ	モ	ヨ	ロ	ヲ	

Because of phonological change, the columns indicated by strikethroughs have no letters in contemporary Japanese, although they were filled in with special letters in classical Japanese. If all the strikethroughs were filled, the chart will contain 50 letters for each of *hiragana* and *katakana*, so the syllabary chart is traditionally called *Gojū-on zu* (chart of 50 sounds). To these should be added the letter ん or ン representing a moraic nasal [N], on the rightmost column.

The "50-sound chart", however, does not exhaust the *hiragana* and *katakana* letters actually employed in Japanese, because the basic consonant sounds (*k, s, t, h*) have variants. The sound represented by the letter *h* is historically related to the sound represented by *p*, and these voiceless obstruents (*k, s, t,* and *p*) have their respective voiced counterparts (*g, z, d,* and *b*). Table 2 shows letters for these consonants followed by five vowels.

Table 2: *Letters for voiced obstruents and bilabial [p]*

transcription	ga	za	da	ba	pa
hiragana	が	ざ	だ	ば	ぱ
katakana	ガ	ザ	ダ	バ	パ
transcription	gi	zi	zi	bi	pi
hiragana	ぎ	じ	ぢ	び	ぴ
katakana	ギ	ジ	ヂ	ビ	ピ
transcription	gu	zu	zu	bu	pu
hiragana	ぐ	ず	づ	ぶ	ぷ
katakana	グ	ズ	ヅ	ブ	プ
transcription	ge	ze	de	be	pe
hiragana	げ	ぜ	で	べ	ぺ
katakana	ゲ	ゼ	デ	ベ	ペ
transcription	go	zo	do	bo	po
hiragana	ご	ぞ	ど	ぼ	ぽ
katakana	ゴ	ゾ	ド	ボ	ポ

It is important to note that Tables 1 and 2 show the conventional letters and alphabetical transcription adopted by the HJLL series; they are not intended to represent the actual pronunciations of Japanese vowels and consonants. For example, among the vowels, the sound represented as "u" is pronounced as [ɯ] with unrounded lips. Consonants may change articulation according to the following vowels. Romanization of these has been controversial with several competing proposals.

There are two Romanization systems widely used in Japan. One known as the Hepburn system is more widely used in public places throughout Japan such as train stations, street signs, as well as in some textbooks for learners of Japanese. This system is ostensibly easier for foreigners familiar with the English spelling system. The *Kunreishiki* (the cabinet ordinance system) is phonemic in nature and is used by many professional linguists. The essential differences between the two Romanization systems center on palatalized and affricate consonants, as shown in Table 3 below by some representative syllables for which two Romanization renditions differ:

Table 3: *Two systems of Romanization*

Hiragana	IPA	Hepburn	Kunreishiki
し	[ʃi]	shi	si
しゃ	[ʃa]	sha	sya
しゅ	[ʃɯ]	shu	syu
しょ	[ʃo]	sho	syo
じ and ぢ	[dʒi]	ji	zi
じゃ	[dʒa]	ja	zya
じゅ	[dʒɯ]	ju	zyu
じょ	[dʒo]	jo	zyo
ち	[tʃi]	chi	ti
ちゃ	[tʃa]	cha	tya
ちゅ	[tʃɯ]	chu	tyu
ちょ	[tʃo]	cho	tyo
つ	[tsɯ]	tsu	tu
づ and ず	[dzɯ]	dzu	zu
ふ	[ɸɯ]	fu	hu

Except for the volumes on Ryukyuan, Ainu, and Japanese dialects, whose phonetics differ from Standard Japanese, HJLL adopts the Kunreishiki system for rendering cited Japanese words and sentences but uses the Hepburn system for rendering conventional forms such as proper nouns and technical linguistic terms in the text and in the translations of examples.

The cited Japanese sentences in HJLL look as below, where the first line transliterates a Japanese sentence in Kunreishiki Romanization, the second line contains interlinear glosses largely following the Leipzig abbreviation convention, and the third line is a free translation of the example sentence.

(1) *Taroo   wa   Ziroo   to   Tookyoo   e   it-te   kutusita   o   kat-ta.*
    Taro   TOP   Jiro   COM   Tokyo   ALL   go-GER   sock   ACC   buy-PST
    'Taro went to Tokyo with Jiro and bought socks.'

The orthographic convention of rendering Japanese is to represent a sentence with an uninterrupted sequence of Sino-Japanese characters and *katakana* or *hiragana* syllabaries without a space for word segmentation, as in 太郎は次郎と東京へ行って靴下を買った for (1). In line with the general rules of Romanization adopted in

books and articles dealing with Japanese, however, HJLL transliterates example sentences by separating word units by spaces. The example in (1) thus has 10 words. Moreover, as in *it-te* (go-GERUNDIVE) and *kat-ta* (buy-PAST) in (1), word-internal morphemes are separated by a hyphen whenever necessary, although this practice is not adopted consistently in all of the HJLL volumes. Special attention should be paid to particles like *wa* (topic), *to* 'with' and *e* 'to, toward', which, in the HJLL representation, are separated from the preceding noun or noun phrase by a space (see section 7.3). Remember that case and other kinds of particles, though spaced, form phrasal units with their preceding nouns.

7.2 Word order

As seen in (1), Japanese is a verb-final, dependent-marking agglutinative language. It is basically an SOV language, which marks the nominal dependent arguments by particles (*wa*, *to*, *e*, and *o* above), and whose predicative component consists of a verbal-stem, a variety of suffixes, auxiliary verbs, and semi-independent predicate extenders pertaining to the speech act of predication (see section 7.6). While a verb is rigidly fixed in sentence final position, the order of subject and object arguments may vary depending on pragmatic factors such as emphasis, background information, and cohesion. Thus, sentence (2a) with the unmarked order below, in principle, may vary in multiple ways as shown by some possibilities in (2b)–(2d).

(2) a. *Taroo   ga   Hanako   ni   Ziroo   o   syookai-si-ta.*
         Taro   NOM   Hanako   DAT   Jiro   ACC   introducing-do-PST
         'Taro introduced Jiro to Hanako.'
     b. *Taroo ga **Ziroo o** Hanako ni syookai-si-ta.*
     c. ***Hanako ni** Taroo ga Ziroo o syookai-si-ta.*
     d. ***Ziroo o** Taroo ga Hanako ni syookai-si-ta.*

Adverbs, likewise, can be rather freely placed, though each type of adverbs has its basic position.

(3) a. ***Saiwainimo** Hanako   ga   gohan   o   tai-te   kure-te   i-ta.*
         luckily   Hanako   NOM   rice   ACC   cook-GER   GIVE-GER   BE-PST
         'Luckily Hanako had done the favor of cooking the rice (for us).'
     b. *Hanako ga **saiwainimo** gohan o tai-te kure-te i-ta.*
     c. *Hanako ga gohan o **saiwainimo** tai-te kure-te i-ta.*

Notice that while the verbal complex in the sentence above is not as tightly organized as a complex involving suffixes, a sentence adverb cannot be placed within the verbal complex, showing that the sequence of *tai-te kure-te i-ta* forms a tighter constituent,

which, however, permits insertion of the topic particle *wa* after each of the gerundive forms. (See section 7.4 below on the nature of gerundive forms in Japanese.)

As the normal position of sentence adverbs is sentence initial, manner and resultative adverbs have an iconically-motivated position, namely before and after the object noun phrase, respectively, as below, though again these adverbs may move around with varying degrees of naturalness:

(4) *Hanako   ga   **isoide**   gohan   o   tai-te   kure-ta.*
    Hanako   NOM   hurriedly   rice   ACC   cook-GER   GIVE-PST
    'Hanako did the favor of cooking the rice hurriedly (for us).'

(5) *Hanako   ga   gohan   o   **yawarakaku**   tai-te   kure-ta.*
    Hanako   NOM   rice   ACC   softly   cook-GER   GIVE-PST
    'Hanako did the favor of cooking the rice soft (for us).'

The fact that an object noun phrase can be easily separated from the verb, as in (2b.d), and that adverbs can freely intervene between an object and a verb, as in (5), has raised the question whether Japanese has a verb phrase consisting of a verb and an object noun phrase as a tightly integrated constituent parallel to the VP in English (cf. **cook hurriedly the rice* – the asterisk marks ungrammatical forms).

7.3 NP structure

Noun phrases, when they occur as arguments or adjuncts, are marked by case particles or postpositions that are placed after their host nouns. Because case markers can be set off by a pause, a filler, or even longer parenthetic material, it is clear that they are unlike declensional affixes in inflectional languages like German or Russian. Their exact status, however, is controversial; some researchers regard them as clitics and others as (non-independent) words.

Elaboration of Japanese noun phrases is done by prenominal modifiers such as a demonstrative, a genitive noun phrase, or an adjective, as below, indicating that Japanese is a consistent head-final language at both nominal and clausal levels.

(6) a. *kono   Taroo   no   kaban*
       this   Taro   GEN   bag
       lit. 'this Taro's bag'

    b. *Taroo   no   kono   kaban*
       Taro   GEN   this   bag
       lit. 'Taro's this bag'

Japanese lacks determiners of the English type that "close off" NP expansion. The literal translations of the Japanese forms above are ungrammatical, indicating that English determiners like demonstratives and genitive noun phrases do not allow further expansion of an NP structure. Also seen above is the possibility that prenominal modifiers can be reordered just like the dependents at the sentence level. The order of prenominal modifiers, however, is regulated by the iconic principle of placing closer to the head noun those modifiers that have a greater contribution in specifying the nature and type of the referent. Thus, descriptive adjectives tend to be placed closer to a head noun than demonstratives and genitive modifiers of non-descriptive types. Interesting is the pattern of genitive modifiers, some of which are more descriptive and are placed closer to the head noun than others. Genitives of the same semantic type, on the other hand, can be freely reordered. Compare:

(7) a. *Yamada-sensei   no   kuroi   kaban*
      Yamada-professor   GEN   black   bag
      'Professor Yamada's black bag'
   b. **kuroi Yamada-sensei no kaban*
      (O.K. with the reading of 'a bag of Professor Yamada who is black')

(8) a. *Yamada-sensei   no   gengogaku   no   koogi*
      Yamada-professor   GEN   linguistics   GEN   lecture
      'Professor Yamada's linguistics lecture'
   b. **gengogaku no Yamada-sensei no koogi*
      (O.K. with the reading of 'a lecture by Professor Yamada of linguistics')

(9) a. *Yamada-sensei   no   kinoo   no   koogi*
      Yamada-professor   GEN   yesterday   GEN   lecture
      lit. 'Professor Yamada's yesterday's lecture' 'Yesterday's lecture by Professor Yamada'
   b. *Kinoo no Yamada-sensei no koogi*

(10) a. *oomori   no   sio-azi   no   raamen*
      big.serving   GEN   salt-tasting   GEN   *ramen*
      lit. 'big-serving salt-tasting *ramen* noodles'
   b. *sio-azi no oomori no raamen*

(11) a. *atui   sio-azi   no   raamen*
      hot   salt-tasting   GEN   *ramen*
      'hot salt-tasting *ramen* noodles'
   b. *sio-azi no atui ramen*

Numeral classifiers (CLFs) pattern together with descriptive modifiers so that they tend to occur closer to a head noun than a possessive genitive phrase.

(12) a. *Taroo   no   san-bon   no   enpitu*
          Taro   GEN   three-CLF   GEN   pencil
          'Taro's three pencils'
      b. **san-bon no Taroo no enpitu*

Numeral classifiers also head an NP, where they play a referential function and where they can be modified by a genitive phrase or an appositive modifier, as in (13a.b). They may also "float" away from the head noun and become adverbial, as in (13c).

(13) a. *Taroo   wa   gakusei   no   **san-nin**   o   mikake-ta.*
          Taro   TOP   student   GEN   three-CLF   ACC   see.by.chance-PST
          'Taro saw three of students by chance.'

      b. *Taroo   wa   gakusei   **san-nin**   o   mikake-ta.*
          Taro   TOP   student   three-CLF   ACC   see.by.chance-PST
          lit. 'Taro saw student-threes by chance.'

      c. *Taroo   wa   gakusei   o   **san-nin**   mikake-ta.*
          Taro   TOP   student   ACC   three-CLF   see.by.chance-PST
          'Taro saw students, three (of them), by chance.'

As in many other SOV languages, the so-called relative clauses are also prenominal and are directly placed before their head nouns without the mediation of "relative pronouns" like the English *which* or *who* or "complementizers" like *that*. The predicates in relative clauses are finite, taking a variety of tense and aspect. The subject may be replaced by a genitive modifier. Observe (14a).

(14) a. *Boku   mo   [Taroo   ga/no   kat-ta]   hon   o   kat-ta.*
          I   ADVPART   Taro   NOM/GEN   buy-PST   book   ACC   buy-PST
          'I also bought the book which Taro bought.'

      b. *Boku   mo   [Taroo   ga/no   kat-ta]   no   o   kat-ta.*
          I   ADVPART   Taro   NOM/GEN   buy-PST   NM   ACC   buy-PST
          'I also bought the one which Taro bought.'

The structure used as a modifier in the relative clause construction can also head a noun phrase, where it has a referential function denoting an entity concept evoked by the structure. In Standard Japanese such a structure is marked by the nominalization particle *no*, as in (14b).

7.4 Subject and topic

Some of the sentences above have noun phrases marked by the nominative case particle *ga* and some by the topic marker *wa* for what appear to correspond to the subject noun phrases in the English translations. This possibility of *ga-* and *wa*-marking is seen below.

(15) a. *Yuki*   *ga*   *siro-i.*
          snow   NOM   white-PRS
          'The snow is white.'

      b. *Yuki*   *wa*   *siro-i.*
          snow   TOP   white-PRS
          'Snow is white.'

As the difference in the English translations indicates, these two sentences are different in meaning. Describing the differences between topic and non-topic sentences has been a major challenge for Japanese grammarians and teachers of Japanese alike. The difference in the English translations above, however, is indicative of how these two sentences might differ in meaning. Sentence (15a) describes a state of affairs involving specific snow just witnessed, whereas (15b) is a generic statement about a property of snow unbounded by time. Thus, while (15a) would be uttered only when the witnessed snow is indeed white, (15b) would be construed true even though we know that there are snow piles that are quite dirty.

A similar difference is seen in verbal sentences as well.

(16) a. *Tori* *ga*   *tob-u.*
          bird   NOM   fly-PRS
          'A bird is flying/is about to fly.'

      b. *Tori* *wa*   *tob-u.*
          bird   TOP   fly-PRS
          'Birds fly.'

Non-topic sentences like (15a) and (16a) are often uttered with an exclamation accompanying a sudden discovery of a state of affairs unfolding right in front of one's eyes. The present tense forms (*-i* for adjectives and *-(r)u* for verbs) here anchor the time of this discovery to the speech time. The present tense forms in (15b) and (16b), on the other hand, mark a generic tense associated with a universal statement.

These explanations can perhaps be extended to a time-bound topic sentence seen in (17b) below.

(17) a. *Taroo   ga   hasit-ta.*
       Taro   NOM   run-PST
       'Taro NOM ran.'

    b. *Taroo   wa   hasit-ta.*
       Taro   TOP   run-PST
       'Taro ran.'

That is, while (17a) reports an occurrence of a particular event at a time prior to the speech time, (17b) describes the nature of the topic referent – that Taro was engaged in the running activity – as a universal truth of the referent, but universal only with respect to a specifically bound time marked by the past tense suffix.

    Topics need not be a subject, and indeed any major sentence constituent, including adverbs, may be marked topic in Japanese, as shown below.

(18) a. *Sono   hon   wa   Taroo   ga   yon-de   i-ru.*
       that   book   TOP   Taro   NOM   read-GER   BE-PRS
       'As for that book, Taro is reading (it).'

    b. *Kyoo   wa   tenki   ga   yo-i.*
       today   TOP   weather   NOM   good-PRS
       'As for today, the weather is good.'

    c. *Sonnani   wa   hayaku   wa   hasir-e   na-i.*
       that.way   TOP   quickly   TOP   run-POTEN   NEG-PRS
       'That quickly, (I) cannot run.'

7.4 Complex sentences

As in many Altaic languages, compound sentences in Japanese do not involve a coordinate conjunction like English *and*. Instead, clauses are connected by the use of inflected verb forms, as in (19a) below, where the *-i* ending is glossed in the HJLL series as either INF (infinitive) or ADVL (adverbial) following the Japanese term *ren'yō-kei* for the form. While the *-i* ending in the formation of compound sentences is still used today, especially in writing, the more commonly used contemporary form involves a conjunctive particle *-te* following the *-i* infinitive form, as in (19b) below. In HJLL, this combination is glossed as GER (gerundive), though the relevant Japanese forms do not have the major nominal use of English gerundive forms.

(19) a. *Hana   wa   sak-i,   tori   wa   uta-u.*
       flower   TOP   bloom-INF   bird   TOP   sing-PRS
       'Flowers bloom and birds sing.'

b. *Hana   wa   sa.i-te,   tori   wa   uta-u.*
   flower   TOP   bloom-GER   bird   TOP   sing-PRS
   'Flowers bloom and birds sing.'

Both the *-i* and *-te* forms play important roles in Japanese grammar. They are also used in clause-chaining constructions for serial events (20a), and in complex sentences (20b)–(20d), as well as in numerous compound verbs (and also in many compound nouns) such as *sak-i hokoru* (bloom-INF boast) 'be in full bloom', *sak-i tuzukeru* (bloom-INF continue) 'continue blooming', *sa.i-te iru* (bloom-GER BE) 'is blooming', and *sa.i-te kureru* (bloom-GER GIVE) 'do the favor of blooming (for me/us)'.

(20)   a.   *Taroo   wa   [ok-i/ok.i-te],   [kao   o   ara-i/arat-te],*
            Taro   TOP   rise-INF/rise-GER   face   ACC   wash-INF/wash-GER

            [gohan   o   tabe-ta].
            meal   ACC   eat.PST
            'Taro got up, washed his face, and ate a meal.'

       b.   *Taroo   wa   [sakana   o   tur-i]   ni   it-ta.*
            Taro   TOP   fish   ACC   catch-INF   DAT   go-PST
            'Taro went to catch fish.'

       c.   *Taroo   wa   [aruk-i   nagara]   hon   o   yon-da.*
            Taro   TOP   walk-INF   SIMUL   book   ACC   read-PST
            'Taro read a book while walking.'

       d.   *Taroo   wa   [Hanako   ga   ki-ta   no]   ni   awa-na-katta.*
            Taro   TOP   Hanako   NOM   come-PST   NM   DAT   see-NEG-PST.
            'Taro did not see (her), even though Hanako came.'

(20d) has the nominalized clause marked by the particle *no* followed by the dative *ni*, also seen in (20b) marking the purposive form. Now the *no-ni* sequence has been reanalyzed as a concessive conjunction meaning 'even though'.

7.5 Context dependency

The context dependency of sentence structure in Japanese is much more clearly pronounced than in languages like English. Indeed, it is rare that Japanese sentences express all the arguments of a verb such as a subject (or topic) and an object noun phrase included in the sentences used above for illustrative purposes. A typical dialog would take the following form, where what is inferable from the speech context is not expressed.

(21) a. Speaker A: *Tokorode, Murakami Haruki no   saisin-saku yon-da   ka.*
                   by.the.way Murakami Haruki GEN newest-work read-PST Q
                   'By the way, have (you) read Haruki Murakami's latest work?'

    b. Speaker B: *Un,   moo   yon-da.*
                   uh-hu already read-PST
                   'Uh-hu, (I) already read (it)'.

In (21a) A's utterance is missing a subject noun phrase referring to the addressee, and B's response in (21b) is missing both subject and object noun phrases. In some frameworks, sentences like these are analyzed as containing zero pronouns or as involving a process of "pro drop", which deletes assumed underlying pronouns. This kind of analysis, however, ignores the role of speech context completely and incorporates information contextually available into sentence structure. In an analysis that takes seriously the dialogic relationship between speech context and sentence structure, the expressions in (21) would be considered full sentences as they are.

7.6 Predicative verbal complexes and extenders

Coding or repeating contextually determinable verb phrases, as in (21b), is less offensive than expressing contextually inferable noun phrases presumably because verb phrases have the predication function of assertion, and because they also code a wide range of other types of speech acts and of contextual information pertaining to the predication act. Declarative sentences with plain verbal endings like the one in (21b) are usable as "neutral" expressions in newspaper articles and literary works, where no specific reader is intended. In daily discourse, the plain verbal forms "explicitly" code the speaker's attitude toward the hearer; namely, that the speaker is treating the hearer as his equal or inferior in social standing, determined primarily by age, power, and familiarity. If the addressee were socially superior or if the occasion demanded formality, a polite, addressee honorific form with the suffix *-masu* would be used, as below.

(22) *Hai,   moo   yom-i-masi-ta.*
      yes   already   read-INF-POL-PST
      'Yes, (I have) already read (it).'

The referent honorific forms are used when the speaker wishes to show deference toward the referent of arguments – subject honorific and object honorific (or humbling) forms depending on the type of argument targeted. If (21b) were to be uttered in reference to a social superior, the following would be more appropriate:

(23) *Un,   (Yamada-sensei   wa)   moo   yom-are-ta.*
uh-hu (Yamada-professor TOP) already read-SUB.HON-PST
'Uh-hu, (Professor Yamada has) already read (it).'

This can be combined with the polite ending *-masu*, as below, where the speaker's deference is shown to both the referent of the subject noun phrase and the addressee:

(24) *Hai,   (Yamada-sensei   wa)   moo   yom-are-masi-ta.*
Yes   (Yamada-professor TOP) already read-HON-POL-PST
'Yes, (Professor Yamada has) already read (it).'

As these examples show, Japanese typically employs agglutinative suffixes in the elaboration of verbal meanings associated with a predication act. The equivalents of English auxiliary verbs are either suffixes or formatives connected to verb stems and suffixed forms in varying degrees of tightness. These are hierarchically structured in a manner that expresses progressively more subjective and interpersonal meaning as one moves away from the verb-stem core toward the periphery. For example, in the following sentence a hyphen marks suffixal elements tightly bonded to the preceding form, an equal sign marks a more loosely connected formative, which permits insertion of certain elements such as the topic particle *wa*, and a space sets off those elements that are independent words following a finite predicate form, which may terminate the utterance.

(25) *(Taroo wa) ik-ase-rare-taku=na-katta   rasi-i   mitai   des-u   wa.*
(Taro TOP) go-CAUS PASS DESI=NEG PST CONJEC PRS UNCERT POLCOP PRS SFP
'(Taro) appears to seem to not want to have been forced to go, I tell you.'

The final particle *wa* above encodes the information that the speaker is female. A male speaker would use *yo* or *da yo*, the latter a combination of the plain copula and *yo*, instead of *desu wa* above, or combinations such as *da ze* and *da zo* in rough speech.

Non-declarative Japanese sentences, on the other hand, frequently suppress auxiliary verbs, the copula, and the question particle especially in casual speech, where intonation and tone of voice provide clues in guessing the intended speech act. Casual interrogatives take the form of (26a) with a nominalization marker bearing a rising intonation, marked by the question mark in the transcription, whereas fuller versions have the interrogative particle *ka* or a combination of the polite copula and *ka*, as in (26b).

(26) a. *Moo   kaeru   no?*
already   return   NM
'Going home already?'

b. *Moo kaeru no (desu) ka.*
already return NM (POLCOP) Q
'Going home already?'

Requests are made with the aid of an auxiliary-like "supporting" verb *kureru* 'GIVE (ME THE FAVOR OF...)', its polite form *kudasai*, or its intimate version *tyoodai*, as seen in (27a). Again, these forms are often suppressed in a highly intimate conversation and may result in a form like (27b).

(27) a. *Hayaku kaet-te kure/kudasai/tyoodai.*
soon return-GER GIVE/GIVE.POL/GIVE.INTI
'(Please) come home soon (for me/us).'

b. *Hayaku kaet-te ne.*
soon return-GER SFP
'(Please) come home soon, won't you?'

The use of dependent forms (e.g., the gerundive *-te* form above) as independent sentences is similar to that of subjunctive forms of European languages as independent sentences, as illustrated by the English sentence below.

(28) *If you would give me five thirty-cent stamps.*

Conditionals are used as independent suggestion sentences in Japanese as well. For example, (29a) has a fuller version like (29b) with the copula as a main-clause verb, which can also be suppressed giving rise to the truncated form (29c).

(29) a. *Hayaku kaet-tara?*
quickly return-COND
lit. ' If return quickly.' 'Why don't you go home quickly?'

b. *Hayaku kaet-tara ikaga desu ka.*
quickly return-COND how POLCOP Q
lit. 'How is it if (you) went home quickly?'

c. *Hayaku kaet-tara ikaga?*
quickly return-COND how
'Why don't (you) go home quickly?'

Understanding Japanese utterances requires full recourse to the elements of speech context, such as the nature of the speaker and the hearer and the social relationship between them, the information "in the air" that is readily accessible to the interlocutors, and the formality of the occasion. Indeed, the difficult part of the art of

speaking Japanese is knowing how much to leave out from the utterance and how to infer what is left unsaid.

8 Conclusion

Many of the interesting topics in Japanese grammar introduced above are discussed in great detail in the Lexicon-Word formation handbook and the Syntax volume. The Historical handbook also traces developments of some of the forms and constructions introduced above. The Sociolinguistics volume gives fuller accounts of the sentence variations motivated by context and discourse genre.

References

Chamberlain, Basil H. 1895. *Essay in aid of a grammar and dictionary of the Luchuan language.* Transactions of the Asiatic Society of Japanese, vol. 23 supplement.

Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Miller, Roy A. 1971. *Japanese and the other Altaic languages*. Chicago: University of Chicago Press.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Appendix: List of abbreviations for HJLL

1	first person
2	second person
3	third person
A	agent-like argument of canonical transitive verb
ABL	ablative
ACC	accusative
ACOP	adjectival copula
ADJ	adjective
ADN	adnominal
ADV	adverb(ial(izer))
ADVL	adverbal
ADVPART	adverbial particle
AGR	agreement
AGT	agent
ALL	allative
AN	adjectival noun

ANTIP	antipassive
AP	adverbial particle, adjective phrase
APPL	applicative
ART	article
ASP	aspect
ATTR	attributive
AUX	auxiliary
AUXV	auxiliary verb
C	consonant
CAUS	causative
CLF	classifier
COHORT	cohortative
COM	comitative
COMP	complementizer
COMPL	completive
CONC	concessive
CONCL	conclusive
COND	conditional
CONJEC	conjectural
CONJCT	conjunctive
CONT	continuative
COP	copula
CVB	converb
DAT	dative
D	demonstrative
DECL	declarative
DEF	definite
DEM	demonstrative
DET	determiner
DESI	desiderative
DIST	distal
DISTR	distributive
DO	direct object
DU	dual
DUR	durative
EMPH	emphatic
ERG	ergative
ETOP	emphatic topic
EVID	evidential
EXCL	exclamatory, exclusive
EXPL	expletive
FOC	focus

FUT	future
GEN	genitive
GER	gerund(ive)
H	high (tone or pitch)
HON	honorific
HUM	humble
IMP	imperative
INCL	inclusive
IND	indicative
INDEF	indefinite
INF	infinitive
INS	instrumental
INT	intentional
INTERJEC	interjection
INTI	intimate
INTR	intransitive
IO	indirect object
IRR	irrealis
ITERA	iterative
k-irr	k-irregular (*ka-hen*)
L	low (tone or pitch)
LB	lower bigrade (*shimo nidan*)
LM	lower monograde (*shimo ichidan*)
LOC	locative
MPST	modal past
MVR	mid vowel raising
N	noun
n-irr	n-irregular (*na-hen*)
NCONJ	negative conjectual
NEC	neccessitive
NEG	negative
NM	nominalization marker
NMLZ	nominalization/nominalizer
NMNL	nominal
NOM	nominative
NONPST	nonpast
NP	noun phrase
OBJ	object
OBL	oblique
OPT	optative
P	patient-like argument of canonical transitive verb, preposition, postposition

PART	particle
PASS	passive
PCONJ	present conjectural
PERF	perfective
PL	plural
POL	polite
POLCOP	polite copula
POSS	possessive
POTEN	potential
PP	prepositional/postpositional phrase
PRED	predicative
PRF	perfect
PRS	present
PRES	presumptive
PROG	progressive
PROH	prohibitive
PROV	provisional
PROX	proximal/proximate
PST	past
PSTCONJ	past conjectural
PTCP	participle
PURP	purposive
Q	question/question particle/question marker
QD	quadrigrade (*yodan*)
QUOT	quotative
r-irr	r-irregular (*ra-hen*)
REAL	realis
RECP	reciprocal
REFL	reflexive
RES	resultative
RESP	respect
S	single argument of canonical intransitive verb, sentence
SBJ	subject
SBJV	subjunctive
SFP	sentence final particle
SG	singular
SIMUL	simultaneous
s-irr	s-irregular (*sa-hen*)
SG	singular
SPON	spontaneous
SPST	simple past
STAT	stative

TOP	topic
TR	transitive
UB	upper bigrade (*kami-nidan*)
UNCERT	uncertain
UM	upper monograde (*kami-ichidan*)
V	verb, vowel
VN	verbal noun
VOC	vocative
VOL	volitional
VP	verb phrase

Languages

ConJ	contemporary Japanese
EMC	Early Middle Chinese
EMJ	Early Middle Japanese
EOJ	Eastern Old Japanese
J-Ch	Japano-Chinese
LMC	Late Middle Chinese
LMJ	Late Middle Japanese
JPN	Japanese
MC	Middle Chinese
MJ	Middle Japanese
MK	Middle Korean
ModJ	Modern Japanese
OC	Old Chinese
OJ	Old Japanese
pJ	proto-Japanese
pK	proto-Korean
SJ	Sino-Japanese
Skt	Sanskrit

Table of contents

Contributors

Kimi Akita completed his Ph.D. at Kobe University in 2009. He taught Cognitive Linguistics at Osaka University before he moved to Nagoya University as Associate Professor in 2015. His research interests include cognitive, typological, and psychological approaches to mimetics and sound symbolism. He has published in journals such as *Cognitive Linguistics*, *Journal of Linguistics*, and *Linguistics*.

Mark Irwin is currently a Professor at the Faculty of Literature and Social Sciences, Yamagata University. His research interests focus not only on loanwords in Japanese, but also on Japanese phonology and the history of the Japanese language. He is the author of *Loanwords in Japanese* (John Benjamins, 2011) and has published widely in journals such as *Journal of East Asian Linguistics*, *Journal of Japanese Linguistics*, and *Phonological Studies*. He is currently working on the *Student Handbook of Japanese* series to be published by Asakura Publishing.

Takane Ito is a Professor in the Department of Language and Information Sciences at the University of Tokyo, Komaba. Her major interest lies in the study of the lexicon and word formation from theoretical and neuro-cognitive perspectives. She co-authored *Go no shikumi to gokeisei* [Mechanisms of word and word formation] (Kenkyusha, 2002) with Yoko Sugioka, and has published papers in journals like *Cognitive Studies*, *English Linguistics*, *Language*, and *NeuroReport*.

Wesley M. Jacobsen is Professor of the Practice of the Japanese Language and Director of the Japanese Language Program at Harvard University. His research is concerned with voice, modality, tense/aspect, and their interaction in Japanese. He is the author of *The Transitive Structure of Events in Japanese* (Kurosio Publishers, 1992) and co-editor of *On Japanese and How to Teach It* (Japan Times, 1990). He has contributed chapters to *Polymorphous Linguistics* (MIT Press, 2005), *Current Issues in the History and Structure of Japanese* (Kurosio Publishers, 2007), and *Nihongo Kyōiku no Atarashii Chihei o Hiraku* (Opening New Horizons in Japanese Pedagogy) (Hituzi Syobo, 2014).

Taro Kageyama (Ph.D., University of Southern California, 1997), co-editor of the HJLL series, is Director-General of the National Institute for Japanese Language and Linguistics (2009–2017). His research interests are in the interaction of morphology, syntax, and semantics. His publications include authored books (*Goi no Kōzō*, 1980, Ichikawa Award; *Bunpō to Gokeisei*, 1993, Kindaichi Award; *Dōshi Imiron* 1996), co-edited books (*Voice and Grammatical Relations* (John Benjamins, 2006), *Transitivity and Valency Alternations* (De Gruyter Mouton)), and journal papers in *Language, Lingua, Yearbook of Morphology*, and *Journal of Japanese Linguistics*.

Hideki Kishimoto (Ph.D., Kobe University, 1991) is Professor of Linguistics at Kobe University, specializing in lexical semantics, morphology, and syntax. He is the author of *Tōgokōzō to bunpō-kankei* [Syntactic structures and grammatical relations] (Kurosio Publishers, 2005) and co-editor of *Goi no imi to bunpō* [Lexical meaning and grammar] (Kurosio Publishers, 2009). He contributed to *The Oxford handbook of Japanese linguistics* (Oxford University Press, 2008) and published articles in *Language, Linguistic Inquiry, Natural Language and Linguistic Theory, Journal of Linguistics, The Linguistic Review, Studies in Language, Journal of East Asian Linguistics*, and *Gengo Kenkyu*.

Hideki Kobayashi is a Professor at the Faculty of Education, Gunma University. He completed his Ph.D. at Osaka University in 2002 and has been working on the structure, meaning, and usage of Sino-Japanese words. His major publications include the authored book *Gendai-nihongo no kango dōmeishi no kenkyū* [A study of Sino-Japanese verbal nouns in contemporary Japanese] (Hituzi Syobo, 2004) and the article "Concerning Sino-Japanese verbs" included in Michiaki Saito and Masahiko Ishii (eds.) *Nihongo goi e no apurōchi* [Approaches to the Japanese vocabulary] (Ōhūsha, 2015).

Tadao Miyamoto is a professor emeritus at Tohoku University. He obtained his second B.A., M.A., and Ph.D. in linguistics from the Department of Linguistics of the University of Victoria, Canada, where he holds an adjunct professor position. He is the author of *Light verb construction in Japanese: The role of verbal nouns* (1999), and is also the co-author (with Joseph F. Kess) of *Japanese psycholinguistics: Classified and annotated bibliography* (1994) and *Japanese mental lexicon: Psycholinguistic studies of kana and kanji processing* (1999), all published by John Benjamins. His research interests include the evolution of language, writing systems, and neuro-imaging studies of language.

Kentaro Nakatani completed his Ph.D. in Linguistics at Harvard University in 2004. He is currently a Professor of Linguistics in the Department of English Literature and Language at Konan University. His main research interest lies in the verb semantics in Japanese, as well as in real-time sentence comprehension and its relation to the use of memory resources. His recent publications include *Predicate Concatenation: A Study of the V-te V Predicate in Japanese* (Kurosio Publishers, 2013) and "A judgment study on aspectual diagnostics in Japanese" (with N. Aoki) in *Japanese/Korean Linguistics* 22 (CSLI Publications, 2015).

Takayasu Namiki is a Professor of Linguistics in the Department of English at Seitoku University and a professor emeritus at Ibaraki University. His research interests focus on the derivational morphology (word formation) of English and Japanese, specifically on compounding from structural and semantic viewpoints. His recent publications include "Morphological variation in Japanese compounds: The case of *hoodai* and the notion of 'compound-specific submeaning'" (*Lingua* 120, 2010). His secondary line of investigation is concerned with Japanese compounds coined from English words.

Yuji Nishiyama completed his Ph.D. at Massachusetts Institute of Technology in 1974. He taught semantics, pragmatics, and the philosophy of language at Keio University and Meikai University. He is Professor Emeritus at both Keio University and Meikai University. His research interests focus not only on the semantics of noun phrases and copular sentences in Japanese, but also on Relevance Theory. He is the author of *Nihongo meishiku no imiron to goyōron* [The Semantics and pragmatics of noun phrases in Japanese] (Hituzi Syobo, 2003) and the editor of *Meishiku no sekai* [On noun phrases] (Hituzi Syobo, 2013).

Naoyuki Ono is a Professor at the Department of Theoretical Linguistics at Tohoku University. His research interest encompasses theoretical and empirical issues in lexical semantics and morphology. He is the author of *Seisei goi imiron* [Generative lexical semantics] (Kurosio Publishers, 2005). His recent publications include "Event structure and the Japanese indirect passive" in J. Pustejovsky et al. (eds.), *Advances in Generative Lexicon Theory* (Springer, 2013) and "Two modes of argument selection in nominals" in *Japanese/Korean Linguistics* 20 (CSLI Publications, 2013).

Michiaki Saito is a Professor at the Graduate School of Arts and Letters, Tohoku University. He completed his Ph.D. at Tohoku University in 1997. His research interest is in Japanese lexicology and word composition. He is the author of *Gendai nihongo no gokōseiron-teki kenkyū* [A Study of the word composition in contemporary Japanese] (Hituzi Syobo, 1992), *Goiron-teki gokōsei-ron* [Lexicological approach to word composition] (Hituzi Syobo, 2004), and *Gokōsei no bunpō-teki sokumen nitsuite no kenkyū* [A study of grammatical aspects of word composition] (Hituzi Syobo, to appear).

Yoko Sugioka is Professor of English and Linguistics at Keio University. She completed her Ph.D. at the University of Chicago in 1984. Since her dissertation, *Interaction of derivational morphology and syntax in Japanese and English* (Garland, 1986), she has mainly worked on morphology and the syntax-lexicon interface. Her other publications include *Go no shikumi to go-keisei* [Mechanism of words and word formation] (with Takane Ito, Kenkyusha, 2002), "Nominalization affixes and the multi-modular nature of word formation" (E. Yuasa et al. (eds.), *Pragmatics and autolexical grammar*, John Benjamins, 2011), and "Incorporation vs. modification in Japanese deverbal compounds" (*Japanese/Korean Linguistics* 10, CSLI Publications, 2002).

Koichi Takezawa received his Ph.D. in Linguistics from the University of Washington in 1987. He is currently Professor in Linguistics and Dean of the College of Japanese Language and Culture at the University of Tsukuba. His research interests center around Japanese syntax and morphology (case marking in particular) and comparative syntax of Japanese and English. He has published articles in anthologies and journals including *Linguistic Inquiry, Linguistic Analysis*, and *English Linguistics*.

Natsuko Tsujimura is Professor in the Department of East Asian Languages and Cultures and Adjunct Professor in the Department of Linguistics at Indiana University, Bloomington. She is the author of *An Introduction to Japanese Linguistics, Third Edition* (Wiley-Blackwell, 2014), and the editor of *Japanese Linguistics: Critical Concepts* (Routledge, 2005) and *The Handbook of Japanese Linguistics* (Blackwell, 1999). Her published articles in journals and book chapters have focused on a wide range of issues involving Japanese verbs as they are relevant to theories in phonology, syntax, semantics, and sociolinguistics.

Satoshi Uehara is Professor of Japanese and linguistics at Tohoku University. He received his Ph.D. in linguistics at the University of Michigan in 1995. His research interests include cognitive linguistics and linguistic typology, with focus on Japanese and East and Southeast Asian languages. His recent publications include "The socio-cultural motivation of referent honorifics in Korean and Japanese" in *Motivation in Grammar and the Lexicon* (John Benjamins, 2011) and "Effects of constituent orders on functional extension patterns of the verbs for give: A Contrastive Study of Thai and Mandarin Chinese" (With Kingkarn Thepkanjana, *Language and Linguistics* 16, 2015).

Kiyo Yamashita is a Professor at the College of Literature, Aoyama Gakuin University. Her research interests are in word formation and meaning in Japanese, with special focus on affixation in Sino-Japanese words. Her recent publications include "Setsuji-sei ji-on keitaiso no zōgo-kinō" in Masa'aki Nomura (ed.), *Gendai nihon kango no tankyū* [Studies on Sino-Japanese affixal morphemes in present-day Japanese] (Tōkyōdō, 2013) and "Sino-Japanese suffixes *-kei* and *-ha*: Focused on derivatives denoting persons" in *Aoyama Journal of Japanese and Japanese Literature* 45, 2015.

Yoko Yumoto is a Professor at the Faculty of Language and Culture, Osaka University. Her publications include "A modular approach to the semantic interpretation of syntactic V-V compounds" (*Journal of Japanese Linguistics* 17, 2001), *Fukugōdōshi to haseidōshi no imi to tōgo* [Syntax and semantics of verbal compounding and prefixation] (Hituzi Syobo, 2005), and "Variation in N-V compound verbs in Japanese" (*Lingua* 120, 2010). She has also co-edited books dealing with lexical semantics and word formation.

Taro Kageyama and Hideki Kishimoto
Introduction

1 Goals and scope of the volume

Words are the pivotal units in the production and comprehension of language, but they are by no means unanalyzable wholes. On the contrary, most words have complex internal structures that are built on the basis of universal as well as language-particular principles of morphology. This handbook presents a comprehensive survey of the lexical phenomena and word formation processes commonly observed in contemporary Japanese, with particular emphasis on their typological characteristics and their interactions with syntax and semantics. The title of the volume could have been "Handbook of Japanese morphology", but the term "morphology" was intentionally avoided in favor of "lexicon and word formation" to make it explicit that the book goes beyond the traditional bipartition between morphology as the realm of words/morphemes and syntax as the realm of phrases/sentences. As stated in the Series Editors' introduction (p. xxix), many of the interesting topics in Japanese grammar cut across the two realms, which are, as it were, bridged by word formation processes. In fact, the majority of chapters contained in this volume show that many of the productive word formation processes in Japanese are not limited narrowly to the domain of morphology but instead function in tandem with syntactic processes, thus giving rise to myriads of interactive phenomena that should interest not only specialists on morphology but also specialists on syntax, semantics, and phonology. This book treats only incidentally phonological issues related to the lexicon, for which the reader is referred to the *Handbook of Japanese phonetics and phonology* (edited by Haruo Kubozono, 2015).

The general orientation of this handbook is not so much theory-oriented as data-oriented. Primary emphasis is placed on the description of basic and common phenomena, especially for the sake of those readers who are not familiar with Japanese, but theoretical issues debated in current theories of the lexicon, word formation, morphology, lexical semantics, and their interaction with syntax are also examined through critically reviewing major works in the previous literature.

The past and ongoing research on lexical matters in Japanese can be conveniently divided into two categories. One is the domestic tradition of philological and descriptive approaches which falls under the general rubric of older *Kokugogaku* ('native language studies' or 'traditional Japanese grammar') or more recent *Nihongogaku* ('Japanese language studies'). The central goal of this tradition is an exhaustive description of raw data based on the collection of actually occurring examples, coupled with their classification and analysis in semantic or quantitative terms. While *Kokugogaku* itself germinated as early as the 18th century in the Edo era, morphological studies that remain influential even today only started in the early

1900s. The other tradition is a theoretical approach that was initiated by linguists working in early generative frameworks, if we leave out an exceptional figure, Bernard Bloch, who analyzed Japanese phonology, morphology, and grammar based on the methodology of American structural linguistics. The fact that early publications of S.-Y. Kuroda, Susumu Kuno, and Masayoshi Shibatani, who are considered the pioneers of theoretical Japanese linguistics in the U.S.A., were addressed to morphosyntactic matters involved in complex predicate formation in syntax is highly suggestive of how intricately Japanese morphology is intertwined with syntax in the whole architecture of grammar. Almost all the chapters in this book indeed make reference, to varying degrees, to syntactic and/or lexical-semantic issues closely tied to the individual morphological topics they tackle.

2 Theoretical significance of Japanese word formation

Words (or lexical items) are often seen as minimal meaningful units of speech. Words are registered in the lexicon, but are by no means static. New words are incessantly added to the lexicon, and words are rendered obsolete as they fall out of use. Natural languages have a number of word formation rules to create complex words. Word formation rules in Japanese are much richer than those in English, and have a wider coverage, because Japanese is a language of an agglutinative type, where morphology plays an important part in forming complex predicates. Japanese has productive lexical processes, such as V-V compounding, that are 'exotic' from Western perspectives, as well as word formation rules, such as N-V compounding, that seem to be governed by some unique conditions.

Modern treatments of word formation, in particular those of morphology and the lexicon, have undergone significant changes through the history of linguistic theory. Various recent theories of word formation have their roots in the early generative approaches of the 1960s and 1970s. In the standard theory (Chomsky 1965), the lexicon played only an impoverished role as a repository for simple lexical items; compound words were derived from sentences by transformational rules (Lees 1960). On the other hand, generative semantics, which enjoyed a heyday from the 1960s until the mid 1970s, abolished the level of deep structure and instead treated even simple words as having complex sources, deriving the verb *kill*, for example, from a complex structure like 'cause to die' (McCawley 1968).

The so-called "lexicalist hypothesis" was launched by Chomsky (1970), where it was proposed that derived nominals (e.g. *destruction*), which have idiosyncratic properties, are created by lexical rules in the lexical component, while gerundive nominals (e.g. *destroying*) are formed by transformation rules in the syntactic component. The lexicalist hypothesis originally postulated that regular syntactic processes do not operate on lexical material, making a clear division of labor between

the lexicon and syntax, such that inflectional morphology, as opposed to derivational morphology, lies outside the domain of the lexicon. Later, a strong lexical hypothesis was proposed to extend the domain of word formation to inflection. Some versions of the lexical hypothesis make a further extension to give a lexical treatment even to the syntactic passive, which was generally assumed to take place in syntax (Wasow 1977; Bresnan 1982). Di Sciullo and Williams (1987) articulated "strong lexicalism", where words are defined as "syntactic atoms" whose internal structure is not accessible to any rule of syntax. In the lexical approaches, various instances of word formation are treated in terms of lexical rules, which are claimed to have distinct properties from syntactic rules. In these lexical theories – generative grammar, HPSG (Head-Driven Phrase Structure Grammar), and LFG (Lexical-Functional Grammar) – various mechanisms are proposed to account for the regularities in word formation processes.

As an agglutinative language, Japanese often creates a gigantic sequence of morphologically complex predicates whose English counterparts are realized by analytic structures such as causative and passive verbs. This gives rise to the issue of how such morphologically complex predicates in Japanese are formed. A number of different treatments are possible with morphologically complex predicates. In the early days of Japanese generative grammar (Kuroda 1965; Kuno 1973; Shibatani 1976), it was customary to derive morphologically complex words from syntactic structures via a transformational rule dubbed "verb raising". On the other hand, the morphological status of complex predicates provided a motive to give them a lexical treatment, as argued by Farmer (1984) and Miyagawa (1980) in a generative framework, by Manning, Sag and Iida (1999) in HPSG, and by Matsumoto (1996) in LFG (see also Sells 1995). Nevertheless, given that complex predicates show properties distinct from single lexical items, many researchers (Sugioka 1986; Kageyama 1982, 1993; Ito and Sugioka 2002; to name just a few) argue that in Japanese, word formation processes apply in the two domains of syntax and the lexicon (apart from the issue pertaining to inflectional morphology). Under this view, regular, productive word formations represent the syntactic type, and irregular, idiosyncratic ones represent the lexical type.

Some recent syntactic approaches assume that complex words can be created entirely in the syntax. Reincarnating the idea of prelexical syntax in generative semantics, Hale and Keyser (1993, 2002) uphold the view that decomposed verb structures are posited at the level of lexical syntax (l-syntax), where verbs are created by following the principles that are essentially the same as those governing the regular syntax. Distributed Morphology (Halle and Marantz 1993) imposes the generative engine on the syntax and claims that morphological words can be created post-syntactically. The recent Minimalist Program does not posit a clear boundary between the lexicon and syntax, and even simple verbs are assumed to have complex syntactic structures (Marantz 1997; Chomsky 1995, 2000).

Japanese is suggestive of several unique kinds of word formation taking place in the realm of the syntax, as well as in the lexicon. Kageyama (1993), for instance, suggests that compound verbs, which constitute single morphological words, can be either syntactic or lexical. For lexical compound verbs, verbs are combined and registered in the lexicon, and syntactic compound verbs have syntactic structure in which one verb is embedded under another. Light verb constructions formed on verbal nouns present another case in point. Verbal nouns have dual properties, in that they can either serve as argument complements to the light verb *suru* 'do' or as predicates in combination with *suru*. Even though argument realization patterns vary, the complex predicates have similar predicative functions regardless of their form, which raises the question of whether the predicate formation takes place at the lexical, syntactic, or LF level.

In the lexicalist hypothesis, the regularities of words should be captured by lexical rules operating in the domain of the lexicon. This view is strongly motivated by the principle of lexical integrity, which states that syntactic operations such as anaphoric binding, movement, and substitution do not affect or alter the internal structures of words (Anderson 1992). The lexical integrity principle is to a large measure observed in Japanese, but there are also some lexical phenomena that do not neatly fit into the standard version of this principle. In Japanese, there is an array of data suggesting that syntactic processes have access to word internal structures, which provides a testing ground for examining how lexical integrity should be treated in the lexicalist hypothesis. Shibatani and Kageyama (1988) show that Japanese has a novel type of post-syntactic word formation, suggesting that word formation can take place after the syntax.

In some current lexical theories (Jackendoff 1990; Grimshaw 1990; Pinker 1989, 2007; Levin and Rappaport Hovav 1995; Pustejovsky 1995), decomposed lexical representations, which are another successor to the prelexical syntactic representations in generative semantics, as well as argument structures, are used to describe linguistically important generalizations over certain well-defined classes of predicates that pattern together. In this respect, Japanese can potentially make a significant contribution because the language has a number of intriguing lexical processes, such as V-V, N-V, and N-N compounding. These are productive word formations processes in Japanese, but they are subject to certain semantic conditions, as there are a number of combinatory restrictions. For instance, Kageyama (1993, 1999) argues that the combinatory patterns of lexical compound verbs are constrained by the Transitivity Harmony Principle, which states that transitive and unergative verbs count as the same type whereas unaccusative verbs form a distinct type. For a compact review of major studies on Japanese word formation, see Kageyama (2014).

Even from the preceding brief discussion of only a few selected instances, it should be apparent that Japanese word formation often shows a number of unique properties, some of which are exotic from the standard Western viewpoint. Although we stressed the interactive relations in Japanese between morphology and syntax, we do not intend to obliterate the boundary between the two realms. There is

definitely an orderly distinction between words/morphemes and phrases/sentences, and the two domains are related to each other by certain rules of word formation that cut across them. A close inspection of the data in Japanese will thus contribute to furthering our understanding of word formation and lexical knowledge, providing a new window into the nature of human language.

3 Organization of the volume

This book comprises nineteen chapters, which are grouped into three parts according to the nature of the topics discussed.

Part I **"Lexicon and vocabulary items"** is intended to lay the foundation for the more advanced discussions in Part II and Part III by explaining fundamental concepts related to the Japanese vocabulary system. Chapter 1 "Vocabulary strata and word formation processes" by Taro Kageyama and Michiaki Saito introduces the basic characteristics of Japanese lexical items and major word formation processes, clarifying the nature of four vocabulary strata – (i) native Japanese, (ii) mimetic or onomatopoeic, (iii) Sino-Japanese, and (iv) foreign – and how they interact with a variety of word formation processes including prefixation, suffixation, compounding, and reduplication. Chapter 2 "Lexical categories" by Hideki Kishimoto and Satoshi Uehara discusses the inventory of Japanese lexical categories from generative and cognitive-functional perspectives, with particular focus on two unique categories, "Verbal Noun" and "Adjectival Noun". Chapter 3 "Sino-Japanese words" by Hideki Kobayashi, Kiyo Yamashita, and Taro Kageyama clarifies the distinct role Sino-Japanese words (Chinese words assimilated to Japanese) play as an independent vocabulary stratum in the Japanese lexicon. Chapter 4 "Mimetics" by Kimi Akita and Natsuko Tsujimura explains how important mimetic words (words that are supposed to represent sounds, manners of action, mental states, etc.) are in the linguistic system of Japanese and how they behave with respect to phonology, semantics, and syntax. Part I closes with Chapter 5 "The morphology of English loanwords" by Mark Irwin, which delineates the morphological properties of borrowings from English, showing how they function not only as nouns but also as verbs, adjectives, and affixes and how they participate in compounding, clipping, and other word formation processes.

Part II **"Morphology and word formation"** covers the major word formation processes in Japanese, revealing not only their morphology but also their semantic and syntactic characteristics. Chapter 6 "Word structure and headedness" by Takayasu Namiki and Taro Kageyama surveys the internal structures of complex words in terms of the notion "head" by dividing them into right-headed, left-headed, double-headed, and headless structures. This chapter also discusses the viability of the distinction between compounds and affixes looking at elements that have undergone grammaticalization or degrammaticalization. Chapter 7 "Noun-compounding and noun-incorporation" by Taro Kageyama shows how the compound words of the

form "Noun + Verb" are organized into a family of mutually related types ranging from tensed compound verbs to tenseless compound verbs through the intermediate category of nonfinite compounds in gerundive and prenominal forms. The discussion on N-V compound verbs is followed by discussion of compound verbs of the form "Verb + Verb", which are used pervasively in Japanese. Chapter 8 "Verb-compounding and verb-incorporation" by Taro Kageyama discusses the distinction between two types of V-V compounds, one lexical and the other syntactic, and highlights their differences in productivity, meaning, and morphological and syntactic behavior. Chapter 9 "Conversion and deverbal compound nouns" by Yoko Yumoto first clarifies the idiosyncratic behavior of verb-to-noun conversion in Japanese and then moves on to probe the semantic complexity of the compound nouns involving deverbal nouns as the head. After these discussions on a variety of compounds comes Chapter 10 "Derivational affixation in the lexicon and syntax" by Yoko Sugioka and Takane Ito, where the distinction between affixes that apply in the lexicon and those that apply in syntactic structure is brought to light with neurological evidence. The subsequent three chapters (Chapters 11 to 13) are devoted to word formation or complex predicate formation in syntactic structure. Chapter 11 "Complex predicates with -*te* gerundive verbs" by Kentaro Nakatani provides an exposition of the complex predicate formation with "auxiliary-like" verbs, Chapter 12 "Light verb constructions with verbal nouns" by Tadao Miyamoto and Hideki Kishimoto presents an overview of the semantic and syntactic properties of complex predicates composed of verbal nouns and *suru* 'do', and Chapter 13 "Inflection" by Koichi Takezawa discusses the syntactic nature of verbal and adjectival inflections. Part II concludes with Chapter 14 "Lexical integrity and the morphology-syntax interface" by Taro Kageyama, where the nature of "lexical integrity" is discussed in relation to the boundary between words and phrases as well as to how word formation interacts with syntax.

Part III "Word classes and syntactic behavior" takes up selected topics showing the close correspondence between semantic classes of words and their syntactic behavior. Chapter 15 "Lexical meaning and temporal aspect" by Wesley M. Jacobsen discusses aspect-based classifications of verbs by comparing the proposals by Haruhiko Kindaichi for Japanese and Zeno Vendler for English. Chapter 16 "Stative and existential/possessive predicates" by Hideki Kishimoto elucidates how various stative predicates in the same semantic class share the same syntactic construction and syntactic behavior. Chapter 17 "Agent nominals" by Naoyuki Ono and Chapter 18 "Complement-taking nouns" by Yuji Nishiyama are addressed to similar classes of nouns that implicate the occurrence of events and actions, the former focusing on human nouns and the latter on non-human nouns, and bring to light interesting correspondences between their meanings and the syntactic constructions they participate in. Part III closes with Chapter 19 "Idioms" by Hideki Kishimoto, where different types of idioms are distinguished by the applicability or non-applicability of a number of diagnostic tests.

Acknowledgments

References

Anderson, Stephen. 1992. *A morphous morphology*. Cambridge: Cambridge University Press.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. Cambridge, MA: MIT Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs and Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. *Minimalist inquiries: The framework*. In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Farmer, Ann K. 1984. *Modularity in syntax: A study of Japanese and English*. Cambridge, MA: MIT Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Hale, Ken and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Ken Hale and Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 53–109. Cambridge, MA: MIT Press.

Hale, Ken and Samuel Jay Keyser. 2002. *Prolegomena to a theory of argument structure*. Cambridge, MA: MIT Press.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and pieces of inflection. In Ken Hale and Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Ito, Takane and Yoko Sugioka. 2002. *Go no shikumi to gokeisei* [Mechanism of words and word formation]. Tokyo: Kenkyusha.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–58.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1999. Word formation. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 297–325. Oxford: Blackwell.

Kageyama, Taro. 2014. Word formation in Japanese. *Oxford bibliographies online*. http://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0194.xml?rskey=ZgBsbv&result=141

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: MIT dissertation. [Published by Garland, 1979.]

Lees, Robert B. 1960. *The grammar of English nominalization*. Bloomington: Indiana University.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, MA: MIT Press.

Manning, Christopher D., Ivan A. Sag and Masayo Iida. 1999. The lexical integrity of Japanese causatives. In Robert D. Levine and Georgia M. Green (eds.), *Studies in contemporary phrase structure grammar*, 39–79. Cambridge: Cambridge University Press.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). 201–225.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structure. In Bill J. Darden, Charles-James N. Bailey and Alice Davison (eds.), *Papers From the Fourth Regional Meeting Chicago Linguistic Society*, 71–80.

Miyagawa, Shigeru. 1980. *Complex verbs and the lexicon*. Tucson AZ: University of Arizona dissertation.

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pinker, Steven. 2007. *The stuff of thought: Language as a window into human nature*. London: Penguin Books.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Sells, Peter. 1995. Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 26. 237–325.

Shibatani, Masayoshi. 1976. Causativization. In Masayoshi Shibatani (ed.), *Syntax and semantics 5: Japanese generative grammar*, 239–94. New York: Academic Press.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar: Postsyntactic compounds in Japanese. *Language* 64. 451–84.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Wasow, Thomas. 1977. Transformations and the lexicon. In Adrian Akmajian, Peter Culicover and Thomas Wasow (eds.), *Formal syntax*, 327–360. New York: Academic Press.

Appendix: List of additional abbreviations for this volume

ANM	animate
ARG	argument
DIR	directional
DP	determiner phrase
DUM.COP	dummy copula
INAN	inanimate
INFL	inflection
MIM	mimetic
PREF	prefix
SUF	suffix
TP	tense phrase
VBLZ	verbalizer

Taro Kageyama and Michiaki Saito

1 Vocabulary strata and word formation processes

1 Introduction

This chapter presents an overview of two concepts basic to analysis of the lexicon of modern Japanese, "vocabulary strata" and "word formation processes", and how they mutually interact in order to provide some background for consideration of individual problems in detail in the following chapters.

The *Nihon kokugo daijiten* (2000–2002), comprising 13 volumes, is the largest dictionary of Japanese ever printed with approximately 600,000 lexical entries and is comparable to *The Oxford English Dictionary* (OED). However, this does not mean that the total number of lexical items in Japanese is 600,000 words. Thanks to the creativity and productivity of language and to the possibility of borrowing words from other languages, an infinite number of additional new words are possible.

The lexicon is composed of a dictionary component that contains all the lexical information traditionally associated with words and morphemes, including morphological, phonological, semantic, syntactic, and stylistic properties, and word formation rules that create new words based on the stored word and morpheme information. Studies on the Japanese lexicon can be divided broadly into descriptive or philological studies, an extension of the Japanese *Kokugogaku* studies tradition, and theoretical research from a generative grammar and cognitive linguistics background. In the former tradition, under the rubric of "word composition", focus is on analyzing the constituents of compounds and derived forms and clarifying how they are combined (Saito 1992a, 2004). However, because the lexicon is treated as a complete system on its own, little attention is paid to mutual relations with other components of the grammar. This tradition began with historical research on affixation and compounding in Old Japanese and ties into research on productivity and word structure in contemporary Japanese. Lately the tradition has linked up with corpus-based lexicology research studying the contexts of word use to illuminate quantitatively sociolinguistic aspects of word use such as gender and age-related differences, dialects, and register (Saito and Ishii (eds.) 2011).

In the theoretical approach, on the other hand, focus is on the abstract "mental lexicon" a native speaker has in his mind and aims to clarify the various constraints on word formation found in Japanese in view of linguistic universals and language-specific realities. The main thrust of research within this tradition has been to locate word formation within the overall architecture of the grammar and to clarify the interactions and interfaces of morphology and syntax, semantics, and phonology. Since the boundary between morphological structure and syntactic structure is often

unclear in an agglutinative language like Japanese, it is necessary to pay attention not only to word formation in the lexicon but also to word formation in syntax. Kageyama (1982), providing an overview of the various word formation processes in Japanese from a 1980s generative grammar viewpoint, was a pioneering work that argued that word formation in Japanese not only takes place within the lexicon but is also related to syntactic structure. Inspired by this work, Sugioka (1986) investigated the interactions between morphology and syntax. As an agglutinative type language, Japanese offers concrete phenomena concerning the great theoretical issue of how to define a "word", or in other words, at what component of the grammar word formation takes place. Shibatani and Kageyama (1988), Kageyama (1993), Matsumoto (1996), and Nishiyama (1998) have gone on to develop their respective theories concerning this question.

In Section 2 of this chapter, we will first explain the vocabulary strata that make up the Japanese lexicon and the characteristics of each stratum. Vocabulary strata are essentially word classes categorized by their historical origin. In the four vocabulary strata generally distinguished, besides vocabulary indigenous to Japan (native Japanese), there are words that entered from Chinese (Sino-Japanese), words borrowed more recently from non-Chinese foreign languages (Foreign), and words that express non-linguistic sounds or cries or vividly express states or actions or physical sensations (Mimetics). In Section 3 we will examine the principal types of word formation in Japanese and their relation to the properties of these strata. It will become clear that the applicability of such word formation processes as compounding, derivation, conversion, reduplication, and clipping varies to a large extent depending on the vocabulary stratum. Finally, in Section 4 we will consider as a topic for future research how to capture the essential properties of these vocabulary strata from a synchronic point of view.

2 Vocabulary strata in contemporary Japanese

In this section we will first describe how the four types of vocabulary strata: native Japanese, Sino-Japanese, Mimetics, and Foreign, are distinguished in the modern Japanese speaker's mental lexicon (Section 2.1) and then explain the principal characteristics of each stratum (Sections 2.2–2.6).

2.1 Distinction of four vocabulary strata

Just as the English lexicon consists of foreign borrowings from French, Latin, Ancient Greek, and other languages overlaid on a basic indigenous vocabulary of Germanic origin, the Japanese lexicon also consists of a number of different vocabulary strata stacked on one another. Word categories based on historical and derivational origins are called *gosyu* or "word types" in traditional Japanese language

studies and four word types are distinguished: a) *wago* (indigenous, native Japanese words), b) *kango* (words entering from Chinese, Sino-Japanese), c) *gairaigo* (words other than *kango* entering from foreign languages since the 16th century, Foreign), and d) *konseigo* (hybrids). As an example of the thinking in traditional Japanese language studies, let us consider the typology presented by Tamamura (1984) in Table 1.

Table 1: Classification of word types in traditional Japanese grammar

Word type	Writing	Examples
1. *Wago* or native words	*hiragana* syllabary *katakana* syllabary Chinese characters	これ *kore* 'this', ある *aru* 'be' ワンワン *wanwan* 'doggy' 山 *yama* 'mountain', 高い *takai* 'tall'
2. *Kango* or Sino-Japanese words	Chinese characters	愛 *ai* 'love', 大学 *daigaku* 'university'
3. *Gairaigo* or Foreign words	*katakana* syllabary Roman alphabet	ミルク *miruku* 'milk', クラス *kurasu* 'class' PC (personal computer)
4. Hybrid words	*hiragana*, *katakana* Chinese characters	あんパン *an pan* (Native-Foreign) 'bean-jam bun', 表玄関 *omote genkan* (Native-S.J.) 'front entrance'

(Adapted from Tamamura 1984: 110)

Three things should be noted about Table 1. First, it associates word type with the four kinds of orthography characteristic to Japanese: *kanji* (Chinese characters), the *hiragana* syllabary, the *katakana* syllabary, and the Roman alphabet. Roughly speaking, Sino-Japanese words are in principle written with Chinese characters, Foreign words in *katakana*, and, among native Japanese words, function words like particles and conjunctions in *hiragana*. However, since each orthography generally has a characteristic perceptual image with *hiragana* having a soft, gentle image, Chinese characters a stiff, formal image, and *katakana* a stylish image, the orthography is sometimes changed depending on the non-linguistic image the writer wishes to convey.

A second point to note about Table 1 is that the category of hybrid words is given the same status as native Japanese, Sino-Japanese, and foreign words. Since the category termed "hybrid" designates words that arise as a result of word formation processes like compounding or derivation, there is no need to posit hybrid as an independent category from the perspective of analyzing modern Japanese.

The third point to note about Table 1 is that mimetics like *wanwan* 'a dog's bark, a doggy' are included in the *wago* category. This type of word is often referred to by the term *onomatope*, 'onomatopoeia', from the French *onomatopée*. Originally, *onomatopée* referred to words mimicking non-linguistic sounds and are sometimes

taken as being "primitive" or "childish" in Western linguistics. In Japanese, however, the range of words covered by the term *onomatope* is much broader than simple onomatopoeia mimicking sounds of nature or the voices of human or animals; it also includes phenomimes (*gitaigo*), which capture the appearance or form of an action or happening or the state or properties of some object and psychomimes (*gijōgo*), which express physical sensations and psychological states. As a term to capture all of these uses and meanings inclusively, the English terms "mimetics" or "mimetic words" are generally used. Since, like the native Japanese stratum, most mimetics are considered to have been present in Japanese from ancient times, before the introduction of Chinese characters, traditional Japanese language studies do not posit an independent category of "mimetics". However, mimetics are not limited to the category of native Japanese words. In fact, there are many types, including some, like *bikkuri* 'surprised', that violate a basic phonological constraint on native words that they should not begin with a voiced obstruent and some, like *seisei-suru* 'feel relieved', in which a Sino-Japanese morpheme is reduplicated. Furthermore, mimetics often differ from other strata in their phonological, morphological, and syntactic behavior (see Chapter 4 [Akita and Tsujimura, this volume]). For these reasons, it is common to treat mimetics as a single class (McCawley 1968; Shibatani 1990; Itô and Mester 1999; Nishio 2002; Irwin 2011; among others).

Defining word types from a purely historical standpoint has many difficulties and is not very useful in analyzing contemporary Japanese. Even though a word may historically originate as a foreign borrowing, in many cases it has been completely assimilated as a fully Japanese word from the perspective of a modern Japanese speaker. For example, *tya* 'tea' is a word that entered Japanese from Chinese and *syake* 'salmon' from Ainu and their word initial palatal consonants [tʃ] and [ʃ] did not originally exist in Japanese. However, these are ancient borrowings and, to a modern Japanese speaker, are fully assimilated into Japanese. For example, *tya* takes the beautific prefix *o-* becoming *o-tya* and when *syake* is compounded as in *beni-zyake* 'red salmon', it undergoes *rendaku* or sequential consonant voicing. The prefixing of *o-* and sequential voicing of consonants are phenomena that normally only occur with native Japanese words.

As seen in the examples of *tya* and *syake*, it is not uncommon for a single given word to simultaneously have characteristics of multiple vocabulary strata. With this caveat, however, we will examine in this section the archetypical characteristics found in the four vocabulary strata: i) *wago* 'Native', ii) *kango* 'Sino-Japanese', iii) *gairaigo* 'Foreign', and iv) Mimetic. Figure 1 illustrates the temporal flow showing the era since which each vocabulary stratum has been a part of Japanese. In contrast to native words and mimetics, or at least a portion of them, that were part of Japanese before the introduction of writing (Chinese characters), Sino-Japanese entered from the middle of the 8th century and foreign words from the 16th century.
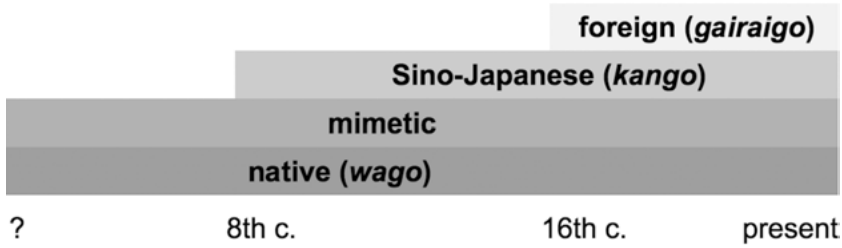
Figure 1: Historical view of vocabulary strata

Besides the historical view, what evidence is there for a distinction of vocabulary strata in the analysis of modern Japanese? The differences a Japanese speaker feels are not just phonological and morphological characteristics of the strata. Germanic-Latinate pairs like *cow-beef* and *pig-pork* are found in English, but Japanese is even richer in having near-synonym sets of items from different vocabulary strata and the difference in strata occasions a predictable difference in the meanings of the synonyms in those sets. Consider the examples in (1).

(1) a. *yado* 'lodging' (native), *ryokan* 'inn' (S-J), *hoteru* 'hotel' (foreign)

   b. *akari* 'illumination' (native), *dentoo* 'electric lamp' (S-J), *raito* 'light' (foreign)

   c. *inu* 'dog' (native), *(keisatu-)ken* '(police-)dog' (S-J), *doggu(-huudo)* 'dog(-food)' (foreign), *wanwan* 'bow wow' (mimetic)

The words given in each set (1a, b, and c) are close in meaning, but they are not completely synonymous. In (1a) the native word *yado* denotes the general concept of lodging, inclusive of Japanese-style lodging (Sino-Japanese *ryokan*) and Western-style (foreign *hoteru*). Similarly, *akari* in (1b) denotes a general concept that includes the S-J *dentoo* and the foreign *raito*. Speaking more generally, when dividing the conceptual level of a noun into "superordinate concept → class concept → subordinate concept", the class concept, which forms the standard for noun categorization, is generally denoted by a simple native word, as in *hi* 'sun', *kaze* 'wind', *ame* 'rain', or *uma* 'horse', *inu* 'dog', *saru* 'monkey' and a subordinate concept is often shown either by a compound native word having the structure (specific difference + class concept) as in *haru-kaze* 'spring breeze', *hama-kaze* 'beach wind', *soyo-kaze* 'refreshing breeze' or by a derived native word like *ko-inu* 'puppy' (Morioka 1970). The characteristic that the native word shows the basic or general concept also applies in the case of verbs and adjectives.

If we posit the four vocabulary strata of Native, Sino-Japanese, Foreign, and Mimetic, we can regard the category of hybrid words, which was problematic for traditional Japanese language studies, as a combination of different strata. The

theoretically possible 12 combinatorial patterns (compounds) of the different strata shown in Table 2 are all attested in reality.

Table 2: Patterns of hybrid words

right-hand left-hand	Native	Sino-Japanese	Foreign	Mimetic
Native	—	**common** *nikai date* 'two-storied'	**common** *mado garasu* 'window pane'	**common** *hara peko* (stomach-empty) 'hungry'
Sino-Japanese	**common** *kiroku yaburi* 'record-breaking'	—	**common** *zyuutaku roon* 'housing loan'	**not common** *nekki munmun* (heat-steamy) 'hot and stifling'
Foreign	**common** *amatyua zumoo* 'amateur *sumo*'	**common** *terebi bangumi* 'TV program'	—	**rare** *pan tira* (panty-showing) 'panty shot'
Mimetic	**common** *garagara hebi* 'rattlesnake'	**common** *tintin densya* 'tinkling train'	**common** *bikkuri man* (surprise-man) 'Bikkuriman'	—

The designations of "common", "not common", and "rare" suggest differences in the productivity of each pattern. Especially notable is the fact that the pattern of a mimetic in the right, head, position and a Sino-Japanese or foreign element in the left, modifier, position lacks productivity. Since, as a general rule, the rightmost (last) element in either syntactic or morphological structure in Japanese is the head, we can make the generalization that mimetics resist becoming heads. This generalization reflects the fact that mimetics archetypically function adverbially (McCawley 1968) and that compounds with adverbial heads are generally rare.

Below we will examine the principal characteristics of the native stratum in Section 2.2, of the Sino-Japanese stratum in Section 2.3, the foreign stratum in Section 2.4, and the mimetic stratum in Section 2.5.

2.2 Native stratum: *Wago*

Words in the native stratum, also called *wago*, are words peculiar to Japanese and form the core of the Japanese lexicon. The *wa* of *wago* originates from ancient Chinese 倭 (*wǒ*; ancient Chinese name for Japan) and the *go* 語 ('word') also comes from Chinese, so the term *wago* itself is from Chinese. The term *Yamato kotoba*

'Yamato language' (*Yamato* being an old name for Japan) is also used to refer to words that are originally Japanese. In the modern orthography, *wago* are generally written in *hiragana* or in a mixture of *kanji* and *hiragana*.

Wago have the following phonological characteristics (Tamamura 1984: 15–23; Labrune 2012: 16; and others).

(2) a. The basic prosodic unit is V (vowel only) or CV (consonant + vowel).

  b. Absence of voiced obstruents /b, d, g, z/ or /r/ in word-initial position.

  c. Absence of word-internal hiatus (consecutive vowels)

  d. Absence of /h/ in word-internal position

  e. Absence of palatalized consonants at word-initial position

  f. Absence of syllabic nasal /N/ (called *hatsu-on* in traditional grammar)

  g. Absence of geminated consonants (called *soku-on* in traditional grammar)

  h. Scarcity of the /e/ vowel

  i. Occurrence of sequential voicing or *rendaku* in nominal compounds

In terms of the length of a single word, measured by the number of mora it has, originally simple words in *wago* fundamentally had one mora, as in *me* 'eye', *te* 'hand', *ha* 'tooth', *ya* 'arrow', and *ke* 'hair', or two moras, as in *asi* 'foot', *koto* 'thing', *aru* 'exist (inanimate)', *hasi* 'bridge'. Words of three or more moras are often complex (compound, derived) words. Since single mora words are felt to be phonetically unstable, there is a strong tendency for single mora words like *ko* 'child', *ta* 'field', *na* 'name', or *ha* 'leaf' to be augmented phonetically or morphologically, as in *kodomo* 'child', *tanbo* 'field', *namae* 'name', or *happa* 'leaf'.

Grammatically, *wago* have the properties necessary to determine the framework of the syntactic structure and the native stratum is the only one that includes all parts of speech including lexical and functional categories. Parts of speech and *wago* examples are given in (3).

(3) a. Noun (N): *hako* 'box', *ki* 'tree', *me* 'eye', *kokoro* 'heart', *sakana* 'fish'

  b. Verb (V): *asob-u* 'play', *su-ru* 'do', *yabur-u* 'tear' (tr.), *yabure-ru* 'tear' (intr.)

  c. Verbal Noun (VN): [compounds] *kasi-kari* 'borrowing and lending', *tati-yomi* 'browsing in a bookshop'

  d. Adjective (A): *maru-i* 'round', *yawaraka-i* 'soft', *kayu-i* 'itchy'

  e. Adjectival Noun (AN): *odayaka-na* 'gentle', *makka-na* 'bright red'

  f. Modal auxiliary (AUX): *rasii* 'it appears', *daroo* 'probably, I think', *yooda* 'it looks like', *sooda* 'I hear'

  g. Adverbial (Adv): *kinoo* 'yesterday', *haya-ku* 'quickly'

  h. Conjunction: *sosite* 'and', *sikasi* 'but', *oyobi* 'and'

  i. Particle: *ga* (nominative), *no* (genitive), *o* (accusative), *kara* 'from', *to* 'with'

Like verbs in English, *wago* verbs show tense through inflectional endings affixed to the verb itself. Japanese does not mark person or number agreement. *Wago* verbs are divided into those with vowel-final stems, such as *tabe-* 'eat' and *koware-* 'break', and those with consonant-final stems, such as *nom-* 'drink' and *wakar-* 'understand'. Using *tabe-* 'eat' as an example of vowel-final stem verbs and *nom-* 'drink' as an example of consonant-final stem verbs, verbal conjugational forms are illustrated in Table 3 (this table is not an exhaustive list of all conjugational forms). For reference, the conjugation of the irregular verb *suru* 'do' is also included in Table 3.

Table 3: Verb inflections

	A. vowel-ending stem (*tabe-* 'eat')	B. consonant-ending stem (*nom-* 'drink')	C. irregular verb *su-* 'do'
1. Past	*tabe ta*	*non da*	*si ta*
2. Non-past	*tabe ru*	*nom u*	*su ru*
3. Infinitive (*ren'yō*)	*tabe*	*nomi*	*si*
4. Gerundive	*tabe te*	*non de*	*si te*
5. Irrealis (*mizen*)	*tabe-* (e.g. *tabe nai* 'eat-not')	*nom a* (e.g. *nom a nai* 'drink-not')	*si* (e.g. *si nai* 'do-not')
6. Hypothetical (*katei*)	*tabe reba*	*nom eba*	*su reba*
7. Cohortative	*tabe yoo*	*nom oo*	*si yoo*
8. Imperative	*tabe ro*	*nom e*	*si ro*

Returning to the list of lexical categories in (3), two categories that do not exist in English or other European languages should be noted: (3c) Verbal Noun (VN) and (3e) Adjectival Noun (AN). Theoretical issues concerning lexical categories will be taken up in Chapter 2 (Kishimoto and Uehara, this volume). The term Verbal Noun is due to Martin (1975) and is widely used in generative and cognitive studies on Japanese morphology. VN is a hybrid category that has the morphology of a noun on one hand but also has verbal functions in that it takes arguments (subject, object) and assigns cases. For example, *tati-yomi* [standing-reading] 'browsing in a bookshop' does not allow direct suffixation of tense endings as in **tati-yom-u* [standing-reading-PRS] or **tati-yon-da* [standing-reading-PST], but must be followed by the verb *suru* in order to express tense. Since *suru* takes the conjugational forms shown in column C of Table 3, *tati-yomi* conjugates as *tati-yomi-su-ru* [standing-reading-do-

PRS] and *tati-yomi-si-ta* [standing-reading-do-PST]. The category VN is thought to have been introduced through Sino-Japanese (*kango*) historically, but, as described later, in contemporary Japanese the category has spread to all vocabulary strata: *wago*, *kango*, foreign, and mimetics. The construction made up of a VN and *suru* is called a "light verb construction" and is discussed in detail in Chapter 12 (Miyamoto and Kishimoto, this volume).

Another hybrid category is that of (3e), Adjectival Noun (AN), which is morphologically a noun but syntactically functions as an adjective. The term Adjectival Noun is due to Martin (1975) and is used by Kageyama (1982, 1993), but Teramura (1982) calls this category *meishiteki-keiyōshi* 'Nominal Adjective'. In traditional Japanese grammar this category is called *keiyōdōshi* 'Adjectival Verb', but this term is based on the conjugational patterns of classical Japanese and is considered inappropriate for modern Japanese. In the same way that verbs and VNs differ, the inflections of normal adjectives and ANs also differ. Japanese adjectives inflect directly, as in Column A of Table 4. As shown in Column B, inflectional endings different from those of adjectives attach to ANs. Unlike verb stems, adjective and AN stems do not vary depending on inflectional endings. A pure adjective has no morphological difference in form between predicative and adnominal use; both forms end in -*i* in the non-past. ANs, however, have the characteristic that they attach -*da* for predicative use and -*na* for adnominal use (rows 2 and 3 in Table 4). See Chapter 2 (Kishimoto and Uehara, this volume), Chapter 13 (Takezawa, this volume), and Chapter 14 (Kageyama, this volume) for detailed discussion on the status of adjectives and ANs.

Table 4: Adjective and AN inflections

	A. Adjective (stem *waka-* 'young')	B. Adjectival Noun (stem *sawayaka-* 'fresh')
1. Past	*waka katta*	*sawayaka datta*
2. Non-past (or Present)	*waka i*	*sawayaka da*
3. adnominal (*rentai*)	*waka i*	*sawayaka na*
4. adverbal (*ren'yō*)	*waka ku*	*sawayaka ni*

Wago make up the basic vocabulary in Japanese and most nouns denoting fundamental concepts like body parts, natural objects, and natural phenomena are *wago*. The fact that most *wago* are basic words of few moras means that a single word designates a broad semantic range. As a result, for instance, the word *a-u* is written with a variety of different *kanji*: 会う, 合う, 遭う, and 逢う. The phenomenon known as *ijidōkun* (expressing the same meaning with different Chinese characters) in traditional Japanese language studies is, from a semantics viewpoint, a case of polysemy where one word expresses several, related, meanings or else is reduced to the phenomenon of homonymy where words with different meanings have the same pronunciation. While on the one hand, there are many cases like 丘 and 岡

(both *oka* 'hill') and 寂しい and 淋しい (both *sabisii* 'lonely') where hardly any clear meaning difference is perceived, there are also cases where the differences in *kanji* are linked to differences in meaning and use. For example, in the case of the *kanji* used for writing *a-u* (会う 'meet (with)', 合う 'match, suit', and 遭う 'encounter'), each can probably be regarded as a separate lexical item with a different meaning. How orthography is related to the linguistic properties of words in Japanese remains to be accounted for (see Nagano and Shimada (2014) for related discussion on the role of *kanji* as representations of lexemes).

Concerning productivity in word formation, *wago* are generally said to be less productive than *kango*, referencing the fact that it is difficult to make long compounds with *wago*. However, several researchers (Tamamura 1975; Saito 1992b; Nishio 2002) have pointed out that this observation is not completely correct. Fresh *wago* compounds are commonly found in specialized fields such as the stock market, for example, where specialized terminology like *taka-domari* [high-stop] 'a stock retains a high price without falling' or *soko-gatai* [bottom-firm] 'the market looks like it could go lower but it does not' are *wago* compounds. Also, the suffix *-sa*, for instance, which nominalizes an adjective or adjectival noun, is widely used not just with *wago* but also with words form the *kango*, foreign, and mimetic strata. Furthermore, this *-sa* has the characteristic that it can even attach to a syntactic clause, as in *sinsya o kaita-sa ni* 'out of desire to buy a new car' (Kageyama 1982; Sugioka 1986; Chapter 10 [Sugioka and Ito, this volume]).

2.3 Sino-Japanese stratum: *Kango*

The difference between the *wago* and *kango* strata in Japanese is similar to the difference between the Germanic stratum and the Latinate (French) stratum in English. In English, the base vocabulary was Germanic and from the 11th through the 15th centuries a vast number of words of Latin origin entered by way of French. Similarly, in Japanese as well, *kango* were piled on a *wago* base stratum from the 8th century on. Just as French was assimilated into English being thought more cultured and refined, in the case of Japanese as well, *kango* entered from China, which had a more advanced culture at the time. Because of this cultural background, *wago*, like Germanic words in English, give an impression of friendliness and ordinariness and belong to a neutral or colloquial register. On the other hand, *kango*, like French- or Latin-origin vocabulary in English, similarly have, on the whole, a stiff, formal atmosphere to them. For example, corresponding to the English pair "buy" and "purchase", Japanese has *kau* and *koonyuu-suru* and corresponding to "build" and "construct", Japanese has *tateru* and *kentiku-suru*.

Kango have the phonological characteristics shown in (4) (cf. Tamamura 1984; Labrune 2012; among others).

(4) a. Occurrence of palatalized consonants

b. Occurrence of voiced obstruments and /r/ at word-initial position

c. Occurrence of moraic nasals /N/ or *hatsu-on* as the coda of a mora

d. Occurrence of long vowels.

e. Occurrence of geminate consonants or *soku-on*

f. Absence of non-geminated /p/

g. Rarity of sequential voicing or *rendaku*

The phonological characteristics of *kango* listed in (4) are not found in the *wago* characteristics listed in (2).

Regarding lexical categories as well, *kango* differ greatly from *wago*. As described above, a defining feature of *wago* verbs and adjectives is that they inflect directly for tense as in *tabe-ru* [eat-PRS] 'eat(s)', *tabe-ta* [eat-PST] 'ate' and *akaru-i* [bright-PRS] 'is bright', *akaru-katta* [bright-PST] 'was bright'. *Kango* alone do not include a conjugable predicate form that can inflect for tense (foreign words and mimetics share this behavior). Since in general, borrowings from foreign languages enter as a noun concept (a kind of quote), even though the word may have been a verb or adjective in the donor language, it can be thought that when it enters Japanese, it is borrowed as a "noun" while retaining its predicating function, becoming the hybrid category VN (Kageyama 1982). VNs are archetypically *kango* and since they require the addition of *suru* to show tense, in traditional Japanese grammar it is customary to call VNs "*suru*-inflected Sino-Japanese nouns". From a linguistic point of view, however, this must be said to be doubly a misnomer. First, as shown by the example of *tati-yomi* in Section 2.2, not all VNs are Sino-Japanese, and second, as shown by examples (5b) and (5c) they need not always be accompanied by *suru*, being capable of taking subjects and objects alone (Shibatani and Kageyama 1988; Kageyama 1993).

(5) a. *Syatyoo*   *wa* *Itaria-sei* *no* *kuruma* *o*
company.president TOP Italian-made GEN car   ACC

koonyuu-si-ta. (with *suru* support)
purchase-do-PST
'The president of the company purchased an Italian car.'

b. *Syatyoo*   *ga* *Itaria-sei* *no* *kuruma* *o*
company.president NOM Italian-made GEN car   ACC

koonyuu *no* *sai* *ni* (without *suru* support)
purchase GEN occasion on
'When the president of the company purchased an Italian car, ...'

c. *Itaria-sei   no   kuruma   o   go-koonyuu   no*
   Italian-made   GEN   car   ACC   HON-purchase   GEN

   okyakusama   wa   (without *suru* support)
   customer   TOP
   'the customer who purchased an Italian car...'

Further, as shown in (5c), a VN can take the subject honorification prefix *go-*.

In addition to noun, VN, and AN, *kango* parts of speech include numeral classifiers used when counting things, such as *-ri/nin* 'persons', *-hiki* 'animals', *-satu* 'volumes', *-mai* 'sheets', and *-dai* 'machines', adverbs like *totuzen* 'suddenly', *issai* 'never do', *kekkyoku* 'in the end', and *motiron* 'of course', and even conjunctions like *naisi* 'from ... to ...' and exclamations (interjections) like *tikusyoo* 'Damn!' and *banzai* 'hooray'.

One morphological characteristic of *kango* is that they come in units of a much wider variety of sizes than *wago*, from bound morphemes that cannot stand alone (root, suffix, prefix), through stand-alone words to a unit that is larger than a normal word but smaller than a phrase, termed "word plus" (Kageyama 2001b; Chapter 14 [Kageyama, this volume]). *Kango* are usually written in *kanji*, but those written with a single *kanji* are mostly bound morphemes. For example, except for a few fixed phrases like *gaku ga aru* 'have an education, be refined' and *gaku o tukeru* 'acquire an education, become refined', 学 *gaku* normally does not stand alone but occurs as part of a compound of two or more characters as in *gakkoo* [*gaku* + *koo*] 'school' or *daigaku* [*dai* + *gaku*] 'college, university'. However, there are also two-character *kango* like *koku-sai* 'international', *sek-kyoku* 'positive', *hon-kaku* 'orthodox', or *bap-pon* 'drastic' that cannot stand alone as free morphemes and require affixation of another *kango* morpheme to form a stand-alone word. Interestingly, when these two-character bound roots form a complex word with another element, they can only occur as the left-hand, modifying, part of the complex as in *kokusai-sei* 'internationalism, internationality', *sekkyoku-teki* 'affirmative', *honkaku-ha* 'fundamentalists', or *bappon-kaikaku* 'drastic reform'. They cannot be moved to the right-hand position to form compounds or derived words as shown by **isi-sekkyoku* 'intention-positive' or **hi-honkaku* 'un-orthodox'.

Regarding the length of *kango* in terms of mora, although there are a few one-more *kango* words like *tya* 'tea' that can be used independently, most are composed of two or more mora. According to Nakano (1973), the vast majority of words used in newspapers are four mora long. Since a single *kanji* is one or two mora long, a *kango* **n** characters long should vary in length from **n** to **2n** moras in length. However, according to Tamamura (1984), of the 2,029 distinct characters in the *Tōyō Kanji Onkunhyō* [List of Sino-Japanese and native readings for tōyō characters], only 488 characters or 24.05% had Sino-Japanese readings of one mora. That is, more than three-fourths of the characters have Sino-Japanese readings of two moras and thus $2 \times 2 = 4$ moras (i.e. two character) words are also common.

Since Chinese tones are not represented in the *kanji* orthography, even cases that are differentiated by the four tones in Chinese end up with the same accent coming into Japanese. In addition, because Japanese has the simple open syllable structure of V or CV, the original syllable structure of Chinese is considerably simplified when borrowed into Japanese. As a result of these phonological constraints, a large number of homophonous *kango* words have developed. Hayashi (ed.) (1982) lists a large number of homophonous words with different meanings culled from the list of homophonous words compiled by the Kokuritsu Kokugo Kenkyūjo (1961), all of which are *kango*. The string *koosyoo* had the greatest number of different meanings, being the pronunciation for 28 homophonous words, all written with different *kanji* compounds, for example 交渉 'negotiation', 口承 'pass on orally', 校章 'school emblem', 公称 'public name' and 公証 'notarization', among others.

Kango also contrast with *wago* in the semantic fields they cover. According to Miyajima (1980), in contrast to *wago*, which are more common in basic vocabulary like natural objects or natural phenomena, in other fields, especially those referring to human intellectual activity, *kango* are preponderant. Morioka (1970) found that words designating superordinate concepts like *doobutu* 'animal', *syokubutu* 'vegetable' or *seibutu* 'animate object', *museibutu* 'inanimate object', rather than class concepts, were nearly all *kango* and that some words designating class concepts, like *zoo* 'elephant', *kirin* 'giraffe' or *ringo* 'apple', *mikan* 'mandarin orange', were also *kango*. However, this does not mean that there are no *wago* denoting abstract concepts. The opposition between *mono*, which shows tangible or intangible objects, and *koto*, which shows events, actions, or states, is the most basic concept for classifying Japanese nouns.

Kango are often said to be more productive in word formation than *wago*. For instance, (6a) is an existing *kango* noun in Japanese and by adding various *kango* elements to it, a longer, more complex word can be created. The examples in (6b) and (6c) do not actually exist, but they are completely natural as Japanese words.

(6)   a.   *danzyo.koyoo.kikai.kintoo.hoo*
              male.female.employment.opportunity.equality.law
              'Equal Employment Opportunity Law'

      b.   [[*danzyo.koyoo.kikai.kintoo.hoo*] *kanren.jikoo*]   (right iteration)
              [[male.female.employment.opportunity.equality.law] related.matters]
              'matters relevant to the Equal Employment Opportunity Law'

      c.   [[*kaisei* [*danzyo.koyoo.kikai.kintoo.hoo*]]   (left iteration)
              [[revision [male.female.employment.opportunity.equality.law]]
              'Revised Equal Employment Opportunity Law'

It is possible to continue to extend (6b) and (6c) even further. However, it is impossible to create such long, complex words with *wago* alone, foreign words alone, or mimetics alone.

It is necessary here to give some attention to the meanings of the terms "word formation ability" or "productivity". In theoretical research in morphology, the terms "productivity" and "creativity" are distinguished and are used in roughly the following meanings (Lyons 1977; Bauer 2001; Kageyama 2010; among others).

(7) a. productivity: a design-feature of the language system

   b. creativity: the language-user's ability to extend the system by means of motivated, but unpredictable, principles of abstraction and comparison

<div align="right">(Lyons 1977: 549; cf. also Bauer 2001: 63)</div>

"Productivity" is the ability to make a new word with a given rule based on a given morphological structure, both of which are determined by the morphology of the language under consideration, and is thought to be one of the abstract mechanisms of human language. "Creativity", on the other hand, is the coining of a neologism taking an existing word as a model, in response to the needs of metaphoric expression or social demands. Since the speaker's expressiveness is motivation for creation of the neologism in this case, people hearing such a coinage for the first time are often impressed by "an unusual way of putting things" or "an interesting, effective expression".

Bearing this distinction in mind, creative formation of neologisms is common with both *kango* and *wago*. Just as modeled on the S-J *koku-sai* lit. 'spanning multiple countries' or 'international', the neologisms *gaku-sai* 'spanning multiple academic fields', *syuu-sai* 'spanning multiple states', and *min-sai* 'exchanges across national borders by multiple citizens groups' have appeared, in *wago* as well a variety of expressions made up of compounds of the pattern *X-irazu* 'the state of X's being unnecessary' can be found, such as *isya-irazu* 'no doctor needed', *neko-irazu* 'no cat needed (rat poison)', *kusuri-irazu* 'no medicine needed', *kasa-irazu* 'no umbrella needed', *tuuyaku-irazu* 'no translation needed', and *pasokon-irazu* 'no personal computer needed'. This sort of neologism formation often stems from a need for naming in social life and deeply involves analogies with existing words.

Compared to these, the long expressions in (6) would not impress a Japanese speaker as being "interesting" or "out of the ordinary" as Japanese expressions. The sort of long, complicated complex words as in (6) are straightforwardly understood, even at first sight, and are interpreted as natural Japanese linguistic structures. Complicated complex words like (6a, b, c) are not arbitrarily created by analogy to existing words, but are the product of a linguistic property inherent in human language called the iterative application of compounding rules. Productivity, the unlimited creation of complex words, is a property of *kango*. Compounds and derived *kango* will be considered in detail in Chapter 3 (Kobayashi, Yamashita, and Kageyama, this volume).

2.4 Foreign stratum: *Gairaigo*

The term *gairaigo* literally means 'words coming from foreign countries', but in the analysis of modern Japanese it refers to borrowings into Japanese from languages other than Chinese since the 16th century, especially in the 19th century, that have a low degree of assimilation into Japanese, in particular, borrowings from western European languages like English, French, German, Italian, and Dutch.

Gairaigo differ greatly from native words in their phonological characteristics: including segments and sequences not found in native words, such as [*wi*, *wo*, *ʃe*, *tʃe*, *tsa*, *tso*, *ti*, *fi*, *fo*, *va*, *vi*, *vo*], allowing geminate voiced obstruents blocked by the phonotactics of native words, as in *baggu* 'bag' and *beddo* 'bed', and allowing /p/ in word-initial position, as in *pasu* 'pass' and *puran* 'plan'. Sequential voicing (*rendaku*) also does not occur with foreign words.

According to Nakano (1973) there are many frequently occurring foreign words with four or five moras, and the fact that the number of moras is greater than in other word types is a feature of foreign words. The reason foreign words are many mora long stems from the fact that the basic syllable structure of Japanese is CV and in order to fit a word from a language like English that allows consonant clusters into the CV pattern, a vowel must be added after each consonant. For example, the word "strike" is only one syllable in English, but when imported into Japanese as *sutoraiku* 'strike in baseball or bowling games' or *sutoraiki* 'labor strike', it becomes five mora long. Clipping often occurs with long foreign words, as it does with long *kango*, but of the words *sutoraiku* and *sutoraiki*, only the latter can be shortened to *suto*; the baseball *sutoraiku* never becomes *suto*. Clipping is discussed in more detail in Section 3.5 and Chapter 5 (Irwin, this volume).

Regarding the semantic fields shown by foreign nouns, as is only natural, they are particularly common for items originating in the West as in clothing (*sukaato* 'skirt', *burausu* 'blouse', *poketto* 'pocket'), machinery (*kamera* 'camera', *terebi* 'television', *razio* 'radio'), materials (*garasu* 'glass', *gomu* 'rubber', *buriki* 'tin') and food (*koohii* 'coffee', *pan* 'bread', *bataa* 'butter') (Miyajima 1980). A great number of these "*katakana* words" have been introduced in specialized fields like science, academics, the arts, economics, and politics and their unnecessary overuse by politicians and the mass media has become a social problem.

Earlier, when discussing word formation with *kango*, we said that creating a neologism in response to expressive need should be distinguished from making long complex words through the iterative application of compounding rules. Research on the lexicon in Japan has overwhelmingly treated the former type of new word creation as "productivity", and Miyajima (1977), Nomura (1984), and Ishino (1992), for example, have pointed to the growth in the number of foreign words in comparison with *wago* and *kango*. In contrast to this, foreign words have almost no rule-governed "productivity" in the sense of creating long complex words. For example, the foreign *tuubaifoo* 'two-by-four [2 × 4]' may appear complex at first

glance, but this is a borrowing from English, including even the preposition, and from a Japanese language view, is simply a single, simple word. When making compounds with this word as a base, it can appear in the left-hand, modifying position as in *tuubaifoo-koohoo* 'two-by-four construction method' and *tuubaifoo-zyuutaku* 'two-by-four housing', but no examples are found with it in the right-hand, head position. Furthermore, it is difficult to expand on it with foreign elements alone; a *kango* is needed in the right-hand, head position.

Finally, concerning parts of speech, foreign words, like *kango*, are basically brought in as nouns. For example, the English verb "sign" has been borrowed into Japanese as *sain*, but it is used either as a noun, as in *sain o morau* 'get a signature', or as a VN, as in *sain-suru* 'sign one's name' (Nishio 2002). As shown by the noun *abekku* 'dating couple' from the French preposition *avec* 'with' and the VN *suruu* *(-suru)* 'ignore, refuse to have anything to do with' from English "through", even a word that is a preposition in the donor language becomes a noun when it enters Japanese. Adjectives become ANs as shown by "elegant" becoming *ereganto-na* and "serious" becoming *siriasu-na*. In addition, phrases in English are borrowed as ANs or VNs, as in "give up" becoming *gibu-appu (-suru)* and "at home" becoming *atto-hoomu (-na)*. Furthermore, foreign words are never used as function words like adverbs or conjunctions. Loanwords from English are more fully discussed in Chapter 5 (Irwin, this volume).

2.5 Mimetic stratum

That words belonging to the mimetic stratum have special morphophonological properties compared with the other strata has long been known. Tamori and Schourup (1999), dividing Japanese mimetics into those with one-mora bases and those with two-mora bases, list the following characteristics (C = consonant, V = vowel, Q = moraic consonant, N = moraic nasal)

(8)   Phonological structure of one-mora base mimetic morphemes

    a.   CV: *hu(to), tu(to)*

    b.   i. CVQ: *kaQ, saQ, haQ, huQ, hoQ*
        ii. CVN: *kaN, koN, tyoN, baN, poN, waN*

    c.   CVV: *kaa, gaa, gyuu, guu, saa, zaa*

    d.   i. CVVQ: *kaaQ, saaQ, suuQ, baaQ, paaQ*
        ii. CVVN: *kaaN, gaaN, kiiN, goon, baaN*

    e.   i. CVQ-CVQ: *kyaQ-kyaQ, kyuQ-kyuQ, kuQ-kuQ*
        ii. CVN-CVN: *kaN-kaN, baN-baN, paN-paN, boN-boN*
        iii. CVV-CVV: *kaa-kaa, gaa-gaa, kyaa-kyaa*

(9)   Phonological structure of two-mora base mimetic morphemes

    a.   CV(C)V: *gaba, gui, soyo, hata, hisi, pui*

    b.   i. CVCVQ: *gusaQ, koroQ, pataQ, basaQ, paraQ*
          ii. CVCV*ri*: *gusari, korori, basari, patari, pokiri*

    c.   CVCVN: *kotuN, goroN, dokiN, bataN, potoN*

    d.   i. CVQCV: *doQka, haQsi, suQku*
          ii. CVNCV: *muNzu, zaNbu*

    e.   i. CVQCV*ri*: *gaQkuri, guQtari, baQsari, baQtari*
          ii. CVNCV*ri*: *geNnari, koNgari, huNwari, boNyari*

Included in these are a wide variety of mimetic words ranging from those that are sound symbolic, mimicking actual sounds, to those that express an occurrence or the manner of an action.

The problem of sound symbolism found with expressions that mimic non-linguistic sounds has long been studied. For example, as seen in the pairs *kataQ* 'something hard falling' and *gataQ* 'something heavy falling', *karakara* 'clatter' and *garagara* 'rattle', *sarasara* 'rustle' and *zarazara* 'rattle', where there is an opposition in the initial consonant between voiceless and voiced, the voiceless alternative is linked with an image of "lightness" and the voiced one with an image of 'heaviness" (Izumi 1976). Tamori and Schourup (1999) say that the suffix *-ri* conveys a "relaxed feeling" or "completion", a word-final geminate consonant conveys "instantaneity", "a sense of speed", or "abrupt ending", a moraic nasal added to the end of the word conveys "resonance", a long vowel in an onomatopoeic word shows a "physically drawn out sound in the world of nature", and repetition signals "repetition or continuation of an action". Similar sound-meaning correspondences in mimetic words are also investigated by Hamano (1998).

Do mimetics have lexical categories? As McCawley (1968) points out, mimetics are basically adverbs, but they have verb-like, adjective-like, and noun-like functions as well. According to Kageyama (2007), as shown in (10), grammatical function changes with a change in the location of accent (capital letters represent high-pitched moras), at least in the Kansai dialect.

(10)   *garagara*

    a.   verbal: *Nodo ga GAragara suru.* 'My throat feels irritated.'

    b.   adverbial: *Iwa ga GAragara to kuzureta.* 'Large boulders came rumbling down.'

    c.   adjectival: *Eigakan wa garagaRA datta.* 'The movie theatre was almost empty.'

    d.   nominal: *Akatyan ni GARAGARA o ageta.* 'I gave the baby a rattle.'

When the first mora is high, as in *GAragara* in (10a) and (10b), the word has a verbal or adverbial function and expresses the dynamic aspect of the situation. When, on the other hand, the last mora is high or the word had a uniform high pitch as in (10c) and (10d), it takes on an adjectival or nominal use and expresses a static aspect. This kind of difference suggests that mimetics do have lexical categories. However, as with foreign words, the mimetic stratum does not have functional categories like conjunction or particle. A comprehensive discussion of mimetics is given in Chapter 4 (Akita and Tsujimura, this volume).

2.6 Quantitative change of vocabulary strata

As a conclusion to the discussion of vocabulary strata, we touch upon the quantitative distribution of the four vocabulary strata. What proportion of the modern Japanese lexicon does each stratum occupy? Also, can any changes in proportion be seen over the relatively short time period of a few decades? There are differences between spoken and written language in the quantitative makeup of the word types; within spoken language further differences arise depending on whether we look at private or public speech; and within written language there are differences depending on whether the materials for analysis are newspapers, magazines, or textbooks. As one example of change over time, we compare two surveys of words used in magazines taken close to 40 years apart and examine how the distribution of the vocabulary strata has changed (note: the mimetic stratum is included in the *wago* count).

Table 5: Change in the distribution of vocabulary strata in magazine articles

		wago	S.-J.	foreign	hybrid
90 magazines surveyed in 1956	type count	36.7%	47.5%	9.8%	6.0%
	token count	53.9%	41.3%	2.9%	1.9%
70 magazines surveyed in 1994	type count	27.9%	36.3%	29.7%	6.1%
	token count	41.6%	45.9%	10.6%	2.0%

Table 5 compares the distribution of vocabulary strata found in two lexical surveys conducted by the Kokuritsu Kokugo Kenkyūjo (1964, 2006). The most striking result in Table 5 is the fact that the proportion of foreign words has roughly tripled over a period of 40 years. In the survey results from 1994, the proportion foreign words occupy of the total number of distinct words (type counts) has outstripped *wago*, confirming with real numbers the intuitive feeling of native speakers that there has been a recent trend toward increasing use of foreign words. However, looking at the results more closely, in both surveys the total of foreign words used as a proportion of total words (token counts) is no more than one-third the proportion of distinct words occupied by foreign words (type counts). In other words, while there is a rich

variety of foreign words, the number of basic, repeatedly used words is small. The increase in the number of foreign words has been called "a deluge of foreign words", but this is no more than a metaphor capturing superficially and impressionistically the phenomenon of a teeming variety of foreign words being used. However, a tendency for a portion of Western words to be absorbed as basic vocabulary does seem to be progressing, as with the use of *toraburu* 'trouble' to comprehensively include the meanings of *momegoto* 'friction between people', *isakai* 'dispute', *huwa* 'discord', *kosyoo* 'out of order', *huguai* 'flaw', and *syoogai* 'obstacle' (Saito and Ishii (eds.) 2011: 44–45). *Wago* show precisely the opposite distribution to foreign words. In both surveys the proportion occupied by *wago* of total words used is higher than the proportion of distinct words. This is because the native stratum includes a large number of basic vocabulary. It is particularly worth noting that in the 1956 survey, *wago* had the highest proportion of total words of the three strata (but were overtaken by *kango* in 1994).

This section has provided an overview of the characteristics of each of the four vocabulary strata. Particularly important for word formation is the type of parts of speech. Given that *wago* included all parts of speech, including both lexical and functional categories, it can be considered the stratum that forms the base of the Japanese lexicon. The stratum with the next largest number of part of speech categories is *kango*. Not having particles and unable to directly inflect for tense, *kango* cannot be said to be completely assimilated into the grammar of Japanese, but the productivity of *kango* in word formation has contributed to giving Japanese its rich vocabulary. Next come foreign words. In terms of parts of speech, foreign words are close to *kango*, but since the foreign stratum does not including adverbs or conjunctions, the degree of assimilation of foreign words is lower than *kango*. Finally come the mimetics which, not having any clear distinction of parts of speech, is weak in terms of syntactic assimilation to Japanese compared to the other strata.

3 Word formation processes in contemporary Japanese

In this section we will describe what sorts of word formation processes Japanese has, giving examples from the four types of vocabulary strata. Looking at the languages of the world, word formation processes may be of the concatenative morphology type, arranging words and morphemes in a linear way, or they may be of the non-concatenative morphology type to which linear alignment does not apply. To illustrate the former type with an example from English, if *high* and *school* are put together in a linear fashion, *high school* results. Addition of *graduation* gives *high school graduation*, forming new words through simple concatenation and the interpretation of the word is derived compositionally from its structural composition.

In contrast to this, in English such phenomena as mutation, which derives a new form through a change of vowel in *sing–sang* or *foot–feet*; reduplication, as in *bye-bye* in which the same form is repeated; and clipping, as in the shortening of *information* to *info*, belong to non-concatenative morphology. In Japanese word formation through vowel changes (mutation) is not common, but otherwise nearly the same processes of word formation are found as in English.

In the following, after outlining the basic morphological properties of Japanese word formation from the standpoints of "word-based" and "morpheme-based" morphology (Section 3.1), we will describe compounding (3.2) and affixation (also called "derivation") (3.3) as examples of concatenative morphology, followed by a brief consideration of conversion (in particular, nominalization of the *ren'yō* or infinitive form of a *wago* verb) (3.4) concerning which opinions regarding categorization as concatenative or non-contatenative morphology is divided. Next will be a brief look at processes belonging to non-concatenative morphology, including reduplication (3.5), clipping (3.6), blending (3.7), acronym and initialism (3.8), and, as a special example, transposition (3.9).

3.1 Words and bound morphemes in word formation

In general theories of morphology, two major approaches have been proposed as to what kind of linguistic unit constitutes the base for word formation. Does word formation operate on words or on morphemes? In the "word-based" morphology theory of Aronoff (1976), regular, rule-based word formation applies at the level of the word, an independently occurring unit, as exemplified by the English compound word *basketball* (< independent word *basket* + independent word *ball*) and the derived word *quickly* (< independent word *quick* + suffix *-ly*), and complex words composed of non-independent morphemes are not formed by word formation rules but have an internal structure that is captured through word analysis, as in *re-mit*, *sub-mit*, and *trans-mit*. The motivation for this approach lies in the observation that compounds in English are basically formed of a combination of two independent words. The theory of word-based morphology contrasts with the theory of "morpheme-based" morphology (Halle 1973), an approach stemming from structural linguistics, where word formation may operate over bound morphemes, so that the *X-mit* words mentioned above, for example, are handled by a rule of word formation that attaches the prefix *X* to the base *mit* (meaning 'send'). In Japanese, it is not only full-fledged free-standing words that make compounds, but non-free morphemes ("bound roots") also participate highly productively in the formation of complex words, as will be shown below. A full description of the Japanese word formation system will thus call for an integration of the merits of both approaches.

Japanese words are traditionally classified into two groups, inflecting words (*yōgen* lit. 'predicate words') and non-inflecting words (*taigen* lit. 'nominal words').

Non-inflecting words, in particular nouns, are generally free-standing units and participate in word formation in their dictionary forms. Thus two independent nouns, *pasokon* 'personal computer' and *desuku* 'desk', both in the Foreign stratum, are combined into the compound *pasokon-desuku* 'PC desk' without any morphological change. *Wago* compound nouns also appear at first glance to be composed of two independent words. For example, the compound *ame-onna* 'a woman who is supposed to bring rainy weather whenever she comes' appears to be a simple combination of the two words *ame* 'rain' and *onna* 'woman'. However, in the compounds *ama-asi* 'the way it rains' from *ame* 'rain' + *asi* 'leg', *saka-ya* 'liquor store' from *sake* 'liquor' + *ya* 'shop', and *ko-kage* 'shade of a tree' from *ki* 'tree' + *kage* 'shade', the final vowels of the first elements, *-e* and *-i*, have changed to *-a* and *-o*, respectively, and the variant forms *ama-*, *saka-*, and *ko-* only appear as the first members of compound words. There are also many cases where a phonological change in the initial segment of the second element occurs, for example insertion of *-s-* in *haru-same* 'spring rains' from *haru* 'spring' + *ame* 'rain' and voicing of an unvoiced consonant (so-called *rendaku*) as in *tabi-datu* from *tabi* 'journey' + *tatu* 'set out'. Again the variants *same* for 'rain' and *datu* for 'set out' do not function as independent words. There are many proposed explanations for these phonological alternations, including historical studies, but whatever the explanation, these phonological phenomena serve the role of giving the two joined elements of the compound a phonetic unity. Since such phonological changes generally arise at the level of tightly bound morphemes, it can be thought that *wago* compounds that appear to be two words joined together are actually not "word + word" but are really "bound morpheme (root) + bound morpheme (root)".

The bound status of compounding elements is even more apparent in inflecting words, i.e. verbs and adjectives in the *wago* stratum. It is a general rule of Japanese morphology that native verbs and adjectives undergo compounding, derivation, and other kinds of word formation in specific bound forms that are stripped of inflectional endings. When adjectives occur in the left-hand position of complex words, they appear as bare "stems", such as *waka-* 'young' in the compound *waka-mono* [young$_{STEM}$-person] 'youngster' or *atataka-* 'warm' in the derived noun *atataka-mi* [warm$_{STEM}$-SUF] 'warmth' (adjectival stems are generally bound and can stand alone only in a special use as exclamations in colloquial speech, as in *Waka!* 'Oh, he/she is so young!'). The bases of verbs for word formation, on the other hand, are twofold. When they appear as the left-hand member of a compound, the infinitive (or *ren'yō*) forms serve as bases, such as *tabe* 'eat' in *tabe-mono* [eat$_{INF}$-thing] 'food' and *nomi* 'drink' in *nomi-mono* [drink$_{INF}$-thing] 'drinks'. Most verb infinitives are also bound except for cases that are lexicalized as nouns, as in *tutumi* 'package' (< *tutum-u* 'to wrap'). When verbs undergo suffixation, on the other hand, the bare stems serve as the base, as in *nom-* 'drink' for *nom-ase(ru)* [drink$_{STEM}$-causative suf. (PRS)] 'make someone drink' or *yorokob-* 'be delighted' for *yorokob-asi(i)* [be.delighted$_{STEM}$-SUF (PRS)] 'delightful'. It is not known why verb infinitives are exploited in the left-

hand position of compounds. Incidentally, inflectional or conjugational endings are invariably attached to stems for both verbs and adjectives, as in *nom-u* 'drink' [drink$_{STEM}$-PRS] (< *nom-* + *-ru*) and *yasasi-i* 'be kind' [kind$_{STEM}$-PRS].

The combining forms of verbs and adjectives are summarized in Table 6, and the possible combinations of bound and free morphemes available in Japanese complex words are surveyed in Table 7.

Table 6: Bases of verbs and adjectives for word formation

	verbs as bases	adjectives as bases
compounding	infinitive (or *ren'yō*) + X	stem + X
suffixation	stem + suffix	stem + suffix

Table 7: Combinations of bound and free morphemes

	compounding	derivation
free + free	*pasokon desuku* [PC$_N$-desk$_N$]	—
bound + bound	*ama gasa* [rain$_N$-shade$_N$] 'umbrella', *rai niti* [come$_{VN}$ (S-J))-Japan$_N$ (S-J)] 'visit to Japan' (cf. Section 3.2)	*saka ya* [liquor$_N$-SUF] 'liquor shop', *atataka mi* [warm$_{A\text{-}STEM}$-SUF] 'warmth'
free + bound	*mizu maki* [water$_N$-sprinkle$_{V\text{-}INF}$] 'watering'	*sinsi rasi(i)* [gentleman$_N$-SUF (PRS)] 'gentlemanly'
bound + free	*nomi mono* [drink$_{V\text{-}INF}$-thing$_N$] 'drinks', *huru hon* [old$_{STEM}$-book$_N$] 'used book'	*ma minami* [PREF-south$_N$] 'due south'

Although the stems and infinitives are, in principle, morphologically bound, when we show their lexical categories they are simply labeled V and A, respectively, and their morphological status will be tagged only when necessary, as in V$_{INF}$ and A$_{STEM}$. Furthermore, inflectional endings will be sometimes shown in parentheses, as in *nom(u)* 'drink' and *yasasi(i)* 'be kind', to represent their inflectional status.

3.2 Compounding

Compounds are often characterized as a word composed of more than one word (or free morpheme), but Japanese exhibits compounding of two bound morphemes as well. A productive example is found in two-character Sino-Japanese VNs of the pattern "verbal morpheme + nominal morpheme". Such VNs range from those that are highly productively expandable to those of low productivity, but as one example of high productivity, we can point to the pattern *rai-X* meaning 'come to some country

or city' (Sibata 1978). As long as there is a place name that can be represented by a single character "X", the possibilities are unlimited. All of these can be used with *suru* 'do' suffixed.

(11) a. Cities: *rai-han* 来阪 'come to Osaka', *rai-sin* 来神 'come to Kobe or Kanagawa', *rai-yoo* 来葉 'come to Chiba', *rai-huku* 来福 'come to Fukuoka or Fukushima', *rai-satu* 来札 'come to Sapporo'

b. Countries and states: *rai-niti* 来日 'come to Japan', *rai-bei* 来米 'come to the U.S.A.', *rai-hutu* 来仏 'come to France', *rai-ei* 来英 'come to England', *rai-goo* 来豪 'come to Australia', *rai-in* 来印 'come to India', *rai-hu* 来布 'come to Hawai'i', *rai-ka* 来加 'come to California'

On the other hand, two-character words of the pattern "*satu* 'kill' + noun (target of killing)" are not very productive. Words of this pattern include *sak-kin* [kill-microbe] 'sterilize', *satu-zin* [kill-person] 'murder', *sas-so* [kill-rat], and *sat-tyuu* [kill-insect], but of these only *sak-kin* can function as a verb by attaching *suru* as in *sakkin-suru* 'sterilize'. Since *satuzin* 'murder' can be used in a collocation like *satuzin o okasu* 'commit murder', but not with *suru* as *satuzin-suru, it is considered to be a noun, not a VN. Furthermore, *sasso* and *sattyuu* are bound roots and in order to stand as a word require support from a following element as in *sasso-zai* [kill.rat-drug] 'rat poison' or *sattyuu-ki* [kill.insect-container] 'bug zapper'. In spite of the limited productivity, it is clear that VNs of the form *satu-X* involve a morphologically transparent structure comprising two bound morphemes. The internal structure of *kango* and their productivity will be further discussed in Chapter 3 (Kobayashi, Yamashita, and Kageyama, this volume).

As we saw above, both independent words and non-independent bound roots can participate to varying degrees in word formation in Japanese. Bearing this fact in mind, let us examine some concrete examples of compounding. In compounds of the pattern "X-Y", the part of speech of the compound as a whole is archetypically determined by lexical category of the element that comes after. For example, in *hude-bako* 'pencil box' (N-N), *nage-nawa* 'lasso' (V-N), and *naga-banasi* 'long talk' (A-N), the categories of the first element are all different (N, V, and A), but, since the second elements are all N, the whole combination is a compound noun. Rephrasing, the right-hand element is the head that determines the lexical category of the whole compound (there are also quite a few examples that do not follow this "right-hand head rule" and they will be discussed in Chapter 6 [Namiki and Kageyama, this volume]).

The examples given in (12) through (14) all have right-hand heads that determine the category of the compound and transparent internal semantic relations.

(12) Compound nouns

    a.   Native: N-N *hude-bako* 'pencil box', V$_{INF}$-N *nage-nawa* [throw-rope] 'lasso' A$_{STEM}$-N *naga-banasi* 'long talk'

    b.   Sino-Japanese: *zen-zitu* [previous$_{BOUND}$-day$_{BOUND}$] 'previous day', *koo-en* [public$_{BOUND}$-garden$_{BOUND}$] 'park', *denki-zidoosya* [electric-car], *yuumei-haiyuu* [famous-actor], *kokusai-kankei* [international$_{BOUND}$-relation]

    c.   Foreign: *geemu-sentaa* [game-center] 'amusement arcade', *onrain-geemu* 'online game', *sohuto-tenisu* [soft-tennis] 'soft-ball tennis'

    d.   Mimetics: *garagara-pon* [rattle-pop] 'lottery wheel', *don-pati* [bang-flick] 'gun battle'

(13) Compound verbs/VNs

    a.   Native: N-V *nami-ut(u)* [wave-hit] 'to wave, ripple', V$_{INF}$-V *nagare-oti(ru)* [flow-fall] 'to flow down', Adv-V *saki-basir(u)* [ahead-run] 'to jump the gun'

    b.   Sino-Japanese: *syuk-ka (suru)* [break.out$_{BOUND}$-fire$_{BOUND}$ (do)] 'outbreak of fire', *doku-syo (suru)* [read$_{BOUND}$-book$_{BOUND}$ (do)] 'book-reading', *enkaku-soosa (suru)* [remote-control (do)] 'remote control'

    c.   Foreign: *reberu-appu (suru)* [level-up (do)] 'improve'

(14) Compound adjectives/ANs

    a.   Native: N-A *hara-guro(i)* [heart-black] 'black-hearted', A$_{STEM}$-A *hoso-naga(i)* [thin-long] 'long and narrow', V$_{INF}$-A *musi-atu(i)* [steam-hot] 'sultry', N-AN *hude-mame* [pen-be.diligent] 'be a good correspondent'

    b.   Sino-Japanese: N-AN *isi-hakuzyaku* [will-weak] 'weak-willed', VN-AN *kokyuu-konnan* [breathing-hard] 'difficult breathing'

As contrasted with the examples of hybrid compounds combining words of different vocabulary strata given in Table 2, the examples presented in (12), (13), and (14) are intended to illustrate combinations from the same vocabulary stratum, i.e. *wago* + *wago*, *kango* + *kango*, foreign + foreign, and mimetic + mimetic. Considering the vocabulary strata involved, while compound nouns (12) allow all four strata, since mimetics do not function alone as predicates, they do not appear in compounds that do function as predicates, namely compound verbs or VNs (13) or compound adjectives or ANs (14), which are restricted to the *wago*, *kango*, and foreign strata.

    According to Lieber (1992), in English, compound nouns are plentiful but the productivity of verb-based compounds is low. Compared to this, it is a characteristic of Japanese that not only are compound nouns productive but also compound verbs and compound adjectives are also highly productive as categories that can stand

alone as predicates (Kageyama 2009). Moreover, as shown in (13a), a wide variety of categories can come before the V head in compound verbs, such as N in *nami-utu* 'wave, beat fast', V in *nagare-otiru* 'flow down', and Adv in *saki-basiru* 'be impertinent', and as seen in (14a), a variety of categories can also appear before the A head in compound adjectives, such as N in *hara-guroi* 'black-hearted', A (stem) in *hoso-nagai* 'long and slender', and V (inf.) in *musi-atui* 'sultry'. This word formation pattern seen in compound verbs and compound adjectives has been characteristic of Japanese since the time of Old Japanese (Sakakura 1966). Compound verbs of the pattern N-V will be discussed in Chapter 7 (Kageyama, this volume), the V-V pattern in Chapter 8 (Kageyama, this volume), and compound nouns will be covered in Chapter 9 (Yumoto, this volume).

3.3 Affixation

Because Japanese does not have infixes, derivational affixes are limited to prefixes and suffixes. Let us examine the use of prefixes and suffixes with concrete examples from each of the four vocabulary strata. Examples of prefixes include the following.

(15) a. Native prefixes: ***ka****-yowa(i)* [PREF-feeble] 'very feeble', ***bun****-nagur(u)* [PREF-beat] 'knock hard', ***hik****-kukur(u)* [PREF-tie] 'tie up', ***o****-tegami* [PREF-letter] 'your letter', ***ma****-ue* [PREF-above] 'right above', ***nise****-daia* [PREF-diamond] 'false diamond', ***mono****-sizuka* [PREF-quiet] 'very quiet'

  b. Sino-Japanese prefixes: ***zen****-sekai* [PREF-world] 'all the world', ***tyoo****-oogata* [PREF-large.size] 'super large size', ***han****-tokeimawari* [PREF-clockwise] 'counter-clockwise', ***gen****-syatyoo* [PREF-company.president] 'current company president', ***zen****-daitooryoo* [PREF-president] 'ex-president', ***mai****-getuyoobi* [PREF-monday] 'every Monday', ***mu****-kiryoku* [PREF-vigor] 'apathetic', ***moo****-hantai* [PREF-opposition] 'fierce opposition', ***mi****-kaiketu* [PREF-solve] 'unsolved'

  c. Foreign: ***noo****-bura* [no-bra] 'braless', ***mai****-kaa* [my-car] 'private car', ***hai****-tensyon* [high-tension] 'very excited', ***mini****-taiken* [PREF-experience] 'small experience'

  d. Mimetic: ***doka****-yuki* [PREF-snow] 'heavy snowfall', ***zuba****-nuke(ru)* [PREF-stand.out] 'far excel', ***bura****-sagar(u)* [PREF-hang] 'dangle loosely'

As a rule, prefixes add some new meaning to the base but do not change lexical category. Among the *kango*, the negative prefixes *mu-* 'not, no' in *mu-sekinin(-na)* [no-responsibility] 'irresponsible' and *mu-kiryoku(-na)* [no-vigor] 'apathetic' and *hu-* 'un-' in *hu-kigen(-na)* [NEG-temper] 'sullen' are generally assumed to change

categories, but whether or not they are really category-changing prefixes will be examined in Chapter 6 (Namiki and Kageyama, this volume).

Among the important prefixes are *o-* and *go-*, which form honorifics, humbles and beautifics. In principle, these do not attach to mimetics or foreign words and overwhelmingly *o-* attaches to *wago* and *go-* to *kango*. This distinction is particularly clear when forming honorific forms of VNs. In the examples of (16), verbal bases, shown in boldface, are sandwiched by the honorific prefix *o-* or *go-* and the honorific light verb *nasaru* 'do'.

(16) a. Native: *{o/*go}-**dekake** nasaru* 'go out'

     b. Sino-Japanese: *{go/*o}-**syuppatu** nasaru* 'leave'

     c. Foreign: *{*o/*go}-**sutaato** nasaru* 'start'

     d. Mimetic: *{*o/*go}-**urouro** nasaru* 'hang around'

However, attachment to form beautifics of nouns is not as consistent.

(17) a. Native: *{o/*go}-sake* 'Japanese *sake*', *{o/*go}-tegami* 'letter'

     b. Sino-Japanese: *{go/*o}-syu* 'Japanese *sake*', *{o/*go}-tya* 'tea', *{o/*go}-soozi* 'cleaning'

     c. Foreign: *{o/*go}-toire* 'toilet', *{*o/*go}-wain* 'wine'

Next, there are the following kinds of suffixes.

(18) a. Native suffixes: *kanasi-**sa*** [sad-SUF] 'sadness', *atu-**mi*** [thick-SUF] 'thickness', *haru-**mek(u)*** [spring-SUF] 'get springlike', *kodomo-**ppo(i)*** [child-SUF] 'childish', *otoko-**rasi(i)*** [man-SUF] 'manly', *sabisi-**ge*** [lonely-SUF] 'look lonely', *gura-**tuk(u)*** [loose-SUF] 'get loose', *darasi-**na(i)*** [tidiness-SUF] 'slovenly'

     b. Sino-Japanese suffixes: *eiga-**ka*** [movie-SUF] 'make into a movie', *sekai-**teki*** [world-SUF] 'world-wide', *arukari-**sei*** [alkali-SUF] 'alkaline', *roodoo-**sya*** [labor-SUF] 'laborer', *doitu-**zin*** [Germany-SUF] 'German people', *nihon-**huu*** [Japan-SUF] 'Japanese style', *takusii-**dai*** [taxi-SUF] 'taxi fare'

     c. Foreign suffixes: *manga-**tikku*** [manga-tique] 'comical', *kinniku-**man*** [muscle-man] 'Kinnikuman', *Waseda-**izumu*** [Waseda-ism], *basuto-**appu*** [bust-up] 'breast enlargement', *puraisu-**daun*** [price-down] 'discount'

     d. Mimetic suffixes: *huto-**ttyo*** [fat-SUF] 'fatty', *gaki-**ntyo*** [kid-SUF] 'brat'

Very few suffixes belong to the mimetic stratum as in (18d).

There are many suffixes that create derived words in the *wago* stratum. For example, the *wago* adjective-making suffix *-si(i)* attaches to a reduplicated morpheme making such words as *hanabana-si(i)* 'magnificent', *uiui-si(i)* 'innocent', *wakawaka-si(i)* 'youthful', *meme-si(i)* 'effeminate', and *yowayowa-si(i)* 'frail, feminine'. The negative suffix *-na(i)* forms such adjectives as *egetu-na(i)* 'nasty', *tumara-na(i)* 'boring', *obotuka-na(i)* doubtful', *sokke-na(i)* 'curt', *kokoro-na(i)* 'inconsiderate', and *zyosai-na(i)* 'tactful', but in many cases, the form to which it attaches is not an independent word in contemporary Japanese. Among suffixes making verbs is *-bur(u)* which attaches to nouns and adjectives as in *yoiko-bur(u)* 'be a goody two-shoes', *otona-bur(u)* 'put on adult airs', *waka-bur(u)* 'try to act young', and *era-bur(u)* 'act important, put on airs'. The suffix *-tuk(u)* makes verbs quite productively, as in *muka-tuk(u)* 'become irritated', *ira-tuk(u)* 'become impatient', *uwa-tuk(u)* 'be frivolous', *bura-tuk(u)* 'hang around', and *gura-tuk(u)* 'be unsteady', but the base to which it attaches is limited to the mimetic stratum and in many cases the "mimetic + *tuk(u)*" co-exists with a form in which the mimetic is in its reduplicated form followed by *suru*, as shown in *mukamuka-su(ru)* 'be irritated, feel sick', *iraira-su(ru)* 'be irritated', *burabura-su(ru)* 'hang around, loiter', and *guragura-su(ru)* 'be unsteady'. However, the *-tuk(u)* verbs and the reduplicated *su(ru)* verbs have different lexical aspectual properties, the former being momentaneous and the latter durative.

As a suffix that creates innovative denominal verbs, there is *-r*. This suffix attaches to foreign words and mimetics to allow them to function as verbs, but it appears that it once also attached to *wago* in older Japanese as in *kubi-r-u* 'kill by choking the neck (*kubi* 'neck') and *kumo-r-u* 'cloud up (*kumo* 'cloud') (Sakakura 1966).

(19)   a.   Native: *kage* 'shadow' → *[kage-r]-u* 'become cloudy', *yazi* (man's name) → *[yazi-r]-u* 'boo, jeer'

     b.   Foreign: *demo* 'demonstration' → *[demo-r]-u* 'stage a demonstration', *kopii* 'copy' → *[kopi-r]-u* 'make a copy, imitate'

     c.   Sino-Japanese: *ziko* 'accident' → *[ziko-r]-u* 'cause an accident', *kokuhaku* 'confession' → *[koku-r]-u* 'confess one's love'

     d.   Mimetic: *[guzu-r]-u* 'fret', *[paku-r]-u* 'pick, steal'

The case of *-r* attaching to a *wago* as in the examples of (19a) is extremely unproductive now. Rarely it attaches to the name of someone famous to describe some action that person famously took, as *egawa-r-u* 'to carry out one's wishes forcibly without paying heed to others', which was coined in 1979 after an incident involving a baseball player named Egawa. Such coinages are colloquial and transient. Attached to foreign words or mimetics *-r* is fairly productive, but is limited to colloquial or vulgar language.

In the examples of (19), it initially looks as if *-ru* is attached to a noun (e.g. *demo* shortened from "demonstration") turning it into a verb (Miyake 2010; Irwin 2011). However, two morphemes are actually involved; the verbalizing suffix is just *-r* (cf. Tsujimura and Davis 2011) and the *-u* is a variant of the non-past tense inflection (as shown in Table 3, the non-past *-ru* appears after vowel-final verb stems but is realized as *-u* after consonant-final stems). This fact is clear from the inflectional forms, as shown in (20).

(20) a. Present: *demor-u* ← [$_V$ [$_N$ *demo*]-*r*]-*u*

 b. Past: *demot-ta* ← [$_V$ [$_N$ *demo*]-*r*]-*ta*   (regressive assimilation of [r] to [t])

 c. Present negative: *demor-anai* ← [$_V$ [$_N$ *demo*]-*r*]-*anai*

 d. Gerundive: *demot-te* ← [$_V$ [$_N$ *demo*]-*r*]-*te* (regressive assimilation of [r] to [t])

In (20) the underlying forms are on the right side of the arrow and the actual phonetic forms are on the left. If the verb *demoru* were composed by adding *-ru* to the noun *demo*, the past tense should be either *demoru-ta or *demo-ta, but those forms do not exist. Since the correct past tense form is *demotta* as in (20b), the stem of the verb must consist of the noun *demo* with *-r* attached, i.e., [$_V$[$_N$*demo*]-*r*].

With *kango* it is often difficult to tell if there is a compounding element or a suffix. Nomura (1978: 104) gives the examples of *tokubetu-si* 'special city', *mui-son* 'village with no doctor', *heikin-ten* 'average value', and *hensa-ti* 'deviation'. The *si* of *tokubetu-si* and the *son* of *mui-son* are also found in the left position of compounds, as in *si-yakusyo* 'municipal offices' and *son-tyoo* 'village chief', so it is probably possible to regard *tokubetu-si* and *mui-son* as compounds. Comparing *heikin-ten* and *hensa-ti*, however, since the *ten* 'points' of *heikin-ten* can be used alone, the word can be considered a compound, but *ti* does not stand alone and it almost never occurs other than in the last position, as in *suu-ti* 'numerical value', *zettai-ti* 'absolute value', and *saidai-ti* 'maximum value', so it can be taken as a suffix. (The word *ti-iki* 'range (of a function)' is a specialized mathematics term.) Looking at what vocabulary strata a one-character *kango* can co-occur with, if it can have *kango*, *wago*, and foreign words all before it, it can probably be considered a suffix. The reason is that there is a general rule that, when the last element in a compound is a one-character *kango*, the preceding element must also be a *kango*.

(21) a. *-teki* '-like': [S.-J.] *huhen-teki* 'universal', [Native] *watakusi-teki* 'in my view', [Foreign] *huransu-teki* 'French style'

 b. *-ryoo* 'charge': [S.-J.] *zyugyoo-ryoo* 'tuition fee', [Native] *hurikomi-ryoo* 'bank transfer fee', [Foreign] *arubaito-ryoo* 'part-time salary'

 c. *-ti* 'value': [S.-J.] *heikin-ti* 'average value', [Native] *sikii-ti* 'threshold', [Foreign] *buuru-ti* 'Boolean value'

The *-teki*, *-ryoo*, and *-ti* in (21) can all be considered suffixes.

3.4 Conversion

The term "conversion" refers to a change from one lexical category to another without a change in form and is usually used with regard to *wago*. This is because *wago* have inflections and the category of the base form is relatively obvious. In contrast, since the category of the base form in the case of *kango*, foreign words, and mimetics is not obvious, it is difficult to decide morphologically from what category to what category a change may have taken place. For example, used in *yuuzin ni tyuukoku-suru* 'warn a friend', *tyuukoku* is a VN, but in *ooku no tyuukoku* 'many warnings', it is a noun. There is no morphological evidence to show that the noun *tyuukoku* has been produced from the VN *tyuukoku* by conversion. Similarly, with the foreign word *kopii* 'copy', in *kopii-suru* it is a VN and in *nimai no kopii* 'two copies', it is a noun and it is difficult to determine which has derived from which. It is least problematic, then, to consider these sets as being separate lexical entries.

The term "conversion" with regard to Japanese, then, is customarily used almost exclusively to refer to clear changes in lexical category with *wago*. A representative example is the case of a *ren'yō* (or infinitive) form of a verb being nominalized as is (Nishio 1988), but there is also the case of a noun becoming an AN and acquiring adjectival properties (Kageyama 2010). Representative patterns are shown in (22).

(22)   a.   V-to-N conversion
           asobu 'to play' → *asobi* 'playing', *nemuru* 'to sleep' → *nemuri* 'a sleep',
           kangaeru 'to think' → *kangae* 'a thought', *nigiru* 'to grasp' → *o-nigiri*
           'a rice ball'

       b.   N-to-AN conversion
           genkin 'cash' → *genkin-na* 'mercenary', *oogesa* lit. 'a Buddhist priest's big
           stole' → *oogesa-na* 'exaggerated', *masikaku* 'true square' → *masikaku-na*
           'square'

Kageyama (2001a) holds that V-to-N conversion in Japanese basically creates event nouns and that with entity nouns, like *o-nigiri* 'rice balls', there is a tendency either for the meaning to become specialized or for the usage to be restricted. Furthermore, there are many examples of deverbalized nouns from conversion forming compounds like *tati-yomi* (standing-reading) 'browsing in a bookshop' and *yasu-uri* (cheap-selling) 'selling cheaply' as well as of nouns converted from adjectives forming compounds, as in *ki-yowa* (mind-feeble) 'weak-minded'. These will be considered in Chapter 6 (Namiki and Kageyama, this volume) and Chapter 9 (Yumoto, this volume).

3.5 Reduplication

Reduplication is found relatively frequently with *wago*, *kango*, and mimetics and sometimes functions to change the meaning and sometimes to change the syntactic function (Martin 1975: 799–800).

(23) a. Reduplication of *wago* nouns to represent plurality.
ie-ie 'houses', *hito-bito* 'people', *kuni-guni* 'countries', *yama-yama* 'mountains', *hana-bana* 'flowers'

b. Reduplication of *wago* or Sino-Japanese elements to produce adverbials.
iro-iro 'variously', *tune-zune* 'always', *toki-doki* 'sometimes', *san-zan* 'thoroughly', *nen-nen* 'year by year'

c. Reduplication of native *ren'yō* verbs to represent circumstantial activities.
naki-naki 'while crying', *(huhei o) ii-ii* 'while expressing (complaints)', *yasumi-yasumi* 'while taking frequent rests'

d. Reduplication of native adjective or AN stems to produce adverbials.
naga-naga (to) 'at great length', *usu-usu (to)* 'vaguely', *tika-zika* 'in the near future', *iya-iya* 'reluctantly'

e. Reduplication of native A stems, V infinitives, Ns, and mimetics to produce VNs.
uki-uki (suru) '(feel) buoyant', *samu-zamu (to suru)* 'be dreary', *gotu-gotu* (suru) 'be rugged', *noro-noro (suru)* 'be sluggish', *kodomo-kodomo (suru)* 'has all the appearance of a small child'

Children's language is rich in reduplication: *nyan-nyan* 'kitty', *wan-wan* 'doggy', *pon-pon* 'tummy', *buu-buu* 'car', *o-me-me* 'eyes', *hai-hai* 'crawl', *(o-kuti o) kutyu-kutyu* 'gargle'. Of these examples, *o-me-me* from *me* 'eye' and *hai-hai* from *hau* 'crawl' are *wago* and the rest are all mimetics.

A meaning difference may arise depending on whether or not a mimetic is reduplicated, as in (24)

(24) a. *Watasi wa doa o ton to ikkai tatai-ta.*
I TOP door ACC tap PRT once knock-PST
'"Tap", I knocked on the door once.'

b. *Watasi wa doa o ton-ton to nikai tatai-ta.*
I TOP door ACC tap-tap PRT twice knock-PST
'"Tap, tap", I knocked on the door twice."

c. *Watasi wa doa o ton-ton-ton to sankai tatai-ta.*
I TOP door ACC tap-tap-tap PRT three.times knock-PST
'"Tap, tap, tap", I knocked on the door three times.'

In (24) the number of the occurrences of *ton* corresponds to the number of knocks on the door, one knock in (24a), two knocks in (24b), and three knocks in (24c), showing a form-meaning iconicity. This is because in the structure of (24) *ton* mimics an actual physical sound and is treated as a direct quote by use of the quotative particle

to. When, however, a real physical sound is not being mimicked or if the quotative particle is not used, reduplication is limited to one repetition, as shown in (25). The single repetition pattern is common across reduplication of *wago* and *kango* as well.

(25) *Iwa   wa   {gara-gara / *gara / *gara-gara-gara}   kuzure-ta.*
       rock TOP {rumble-rumble / *rumble / *rumble-rumble-rumble} collapse-PST
       'The rocks collapsed rumbling.'

3.6 Blending

Blends (also called "portmanteau words") are formed by combining parts of two separate words to form a single phonetic word expressing a single concept. The name of the Japanese movie beast Godzilla is a well-known blend combining the *go-* of *gorira* 'gorilla' and the *-zira* of *kuzira* 'whale'. Blending can occur with verbs, as in *yabuku* 'tear' from *yaburu* 'tear' and *saku* 'rip', and with adjectives, as in *erokawaii* 'sexy-but-cute' from *eroi* 'erotic' and *kawaii* 'cute', but most blending occurs in noun formation. There is no constraint on the type of word; the *wago*, *kango*, foreign, and mimetic strata are all involved.

(26) a. *dotanba* (S.-J.) 'the last minute' + *kyanseru* (Foreign) 'cancel' → *dota-kayan* 'cancel at the last minute'

      b. *Ginza* (S.-J.) 'Ginza, Tokyo' + *burabura* (Mimetic) 'hang around' → *gin-bura* 'stroll on the Ginza'

      c. *asa* (Native) 'morning' + *syanpuu* (Foreign) 'shampoo' → *asa-syan* 'shampooing in the morning'

      d. *Kokubunji* (place name) + *Tachikawa* (place name) → *Kuni-tachi* (name of a place located between Kokubunji and Tachikawa)

As seen in (26d), blending is also used in the creation of new place names. What is especially interesting about (26d) is the reading of the *kanji*. Taking the first character from each of 国分寺 (Kokubunji) and 立川 (Tachikawa) and combining them gives 国立. If this combination is read as a *kango* (using Sino-Japanese readings), it would be *kokuritu*, which is homophonous with the word meaning 'national'. In order to avoid this accidental homophony, both 国 and 立 are read as *wago* (native Japanese readings), giving *Kunitachi* as the pronunciation.

3.7 Clipping

Clipping is the phenomenon of shortening an existing word phonetically by omitting a part of it and has been extensively studied in phonology (see Kubozono (ed.) 2015).

Clipping in Japanese is frequently found in *wago*, *kango*, and foreign words and is often limited to the colloquial language, although some clipped words have become established in the formal register as well. The examples of (27) are divided by the location of the clipped portion into back-clipping and fore-clipping and include examples from each of the three strata. Square brackets show the slipped portions.

(27) a. Back-clipping
       Native: *sakasa[ma]* 'upside-down', *nigiri[zusi]* 'hand-rolled sushi'
       Sino-Japanese: *tyuugaku[koo]* 'junior high school', *koosoku[dooro]* 'expressway'
       Foreign: *suto[raiki]* 'labor strike', *kone[kusyon]* 'connection', *anime[esyon]* 'animation'

     b. Fore-clipping
       Native: *[tomo]dati* 'friend, buddy', *[wa]sabi* 'wasabi, accentuation'
       Sino-Japanese: *[kei]satu* 'cop', *[ma]yaku* 'drugs', *[kis]saten* 'café'
       Foreign: *[aru]baito* 'part-time job', *[intaa]netto* 'the net', *[puratto]hoomu* 'station platform'

"Back-clipping" and "fore-clipping" are terms referring to where the deletion has taken place, but, actually, what has been left, the "remnant", is more important than what has been deleted. The reason that back-clipping is much more common and is easier to interpret than fore-clipping is that the first element of the word remains and it is easy to guess the original word. In contrast, since only the latter portion of the word remains in fore-clipping, it is more difficult to guess the original word. Words like *dati* 'friend', *satu* 'cop', *yaku* 'drugs' are thus conspicuous and are often limited to the argot of certain groups or to the vulgar style.

In all of the examples in (27), a part of a whole word is omitted. In contrast to English, in which this kind of clipping is found in abundance, one characteristic of Japanese is the formation of compounds from multiple elements, each of which has been clipped. There are only a few such examples in English, like *sci-fi* (science fiction) and *sit-com* (situation comedy), but in Japanese, as shown by the examples in (28) this type of compound formation is extremely frequent with *wago*, *kango*, foreign words, and hybrids of these strata. Moreover, the same process is shared by both common nouns and proper nouns.

(28) *ten[pura]-don[buri]* 'a bowl of rice topped with *tempura*', *tok[ubetu]-kyuu[koo]* 'special express train', *Too[kyoo]-dai[gaku]* 'University of Tokyo', *aru[kooru]-tyuu[doku]* 'alcoholic', *me[e]ru-ado[resu]* or *me[eru]-ado[resu]* 'mail address', *poke[tto]-mon[sutaa]* 'Pokemon', *Kimu[ra] Taku[ya]* (Japanese singer and actor)

The English edition of Wikipedia (http://en.wikipedia.org/wiki/Acronym and initialism) considers this type of Japanese example to belong to acronyms (cf. Section 3.8). However, if this type of clipping were acronyms, only the first character of each word should be targeted. As seen in the examples [*iti*] *man* [*en*] *satu* 'ten thousand yen bill' and [*Wata*]*nabe Sada*[*o*] 'Watanabe Sadao, a Japanese jazz musician', though, there are cases in which a part other than the initial character remains. Therefore, it is more appropriate to treat the examples of (28) as examples of clipping than as acronyms. This type of *kango* ellipsis will be touched on in Chapter 3 (Kobayashi, Yamashita, and Kageyama, this volume).

Incidentally, two character compounds that refer to two cities separately (or together) like *Han-Shin (kan)* '(between) Osaka and Kobe' and *Too-Mei (koosoku-dooro)* 'Tokyo and Nagoya (expressway)' are coordinate compounds (so-called "dvandva" compounds) and can be analyzed as compounds of clipped elements, just like the examples of (28). These are different in nature from a blend like *Kunitati* which names a new third town.

With foreign words, a process opposite to clipping sometimes takes place in which some element is added, creating a compound word that does not exist in the donor language (Kageyama 2002). For example, from "hula" we get *huradansu* 'hula + dance', from "mug" *magukappu* 'mug + cup', "(box) cutter" *kattaanaihu* 'cutter + knife', and "(security) guard" *gaadoman* 'guard + man'. The motivation for creating these "made-in-Japan" English compounds may be that the original English word alone was considered insufficiently clear and so, by creating a new compound with an added head noun, the complete meaning becomes more transparent.

3.8 Acronyms and initialisms

In English morphology, an acronym is a word that is created by combining the initial letters of several words and which is pronounced as an English word, as in *laser* from '*l*ight *a*mplification by *s*timulated *e*mission of *r*adiation', whereas an initialism (also called 'alphabetism') is a word that is created similarly but in which each letter is pronounced as its conventional name, as in *ATM* from 'automated [automatic] teller machine'. In Japanese as well, initialisms (29a) and acronyms (29b) are used in the names of organizations, corporations, and groups.

(29) a. *JR* (Japan Railway Company), *JSPS* (Japan Society for the Promotion of Science), *NHK* (Nihon Hōsō Kyōkai; Japan Broadcasting Cooperation)

      b. *UFO* [juːfo] or [juːho], *VIP* [vippɯ] or [bippɯ], *JAL* [dʑarɯ], *MEXT* [mekɯsɯto] (Ministry of Education, Culture, Sports, Science and Technology)

Even words that are treated as initialisms in English, such as *VIP*, *UFO* and *JAL*, are treated as acronyms in Japanese. The acronym for a meteorological system developed in Japan, *AMeDAS* (Automated Meteorological Data Acquisition System), is pronounced *amedasu* and is a witty word that brings to mind the Kansai dialect form (*Ame dasu*) of *Ame desu* 'It's rain.'

Easily mistaken for initialisms are alphabetic abbreviations of Chinese characters. For example, *K. Y.*, pronounced as *kei wai*, was a popular expression around 2010 and is formed by taking the initial letters from the romanized version of *kuuki o yomenai* (lit. can't read the air) 'blurting out something off topic because of not appreciating the conversational situation' and using them to express the full phrase. This kind of abbreviation, which has been around for a long time as a code intelligible only to insiders, should be distinguished from initialism. In contrast to a true initialism, which can fulfill a syntactic role in a sentence, abbreviations like "K. Y." cannot be words with grammatical properties. They are, instead, similar to English abbreviations like *T.B.A* (to be announced) or *p.t.o* (Please turn over.).

3.9 Transposition

Finally, although it is not a word formation process widely used in Japanese, we take up the phenomenon of creating argot through switching the positions of mora within a word, in a way similar to the inversion type of "verlan" in French such as *garette-ci* from *cigarette*. There are some examples that have become somewhat established, such as *neta* 'material' from *tane* 'seed', *syoba* 'space, location' from *basyo* 'location' and *gasa* 'execution of a search warrant by the police' from *saga(su)* 'search', but many, like *maiuu* 'delicious' from *umai* 'delicious' in the performing world or *yanopii* 'piano' from *piano* 'piano' among jazz musicians (cf. Tateishi 1989), are jargon limited in use to a certain realm or by certain (often anti-social) groups.

4 Summary and future research perspectives

This chapter introduced basic concepts in the Japanese lexicon, presenting an overview of the four kinds of vocabulary strata: Native (*wago*), Sino-Japanese (*kango*), Foreign, and Mimetics and of the characteristics of word formation processes like compounding, derivation, conversion, and reduplication and showing that the four types of strata do not participate equally in all word formation processes. As a conclusion to this chapter we will briefly consider how the relation between vocabulary strata and word formation processes should be captured. The mutual relation between vocabulary strata and word formation processes is an important issue that has not been comprehensively considered from the view of the Japanese lexicon as

a whole. Consideration of this issue will add theoretical support from a modern Japanese synchronic viewpoint to the concept of vocabulary strata that has been treated as a historical product in traditional Japanese *Kokugogaku* studies.

In spite of not knowing the historical origin of individual lexical items, a native speaker, without regard to age or dialect background, consciously or unconsciously apprehends the differences among the vocabulary strata. Vocabulary strata, then, can be thought to be part of the lexical information inherently present in the mental lexicon. So, how can we integrate the concept of vocabulary strata into a synchronic grammar of Japanese. An intriguing solution to this question is proposed in Itô and Mester (1999). However, their proposal is limited in the sense that it holds that the target of the rules and constraints of lexical phonology that are involved in word formation is the "phonological lexicon". Their proposal is that the four vocabulary strata do not exist independent of each other but form a hierarchical structure as shown in (30).

(30)

$$
\begin{array}{l}
\text{Japanese lexicon} \\
\quad\diagup\quad\diagdown \\
\text{Lexicon}^2 \qquad \text{Unassimilated foreign} \\
\quad\diagup\quad\diagdown \\
\text{Lexicon}^1 \qquad \text{Assimilated foreign} \\
\quad\diagup\quad\diagdown \\
\text{Lexicon}^0 \qquad \text{Established loans} \\
\text{Native}
\end{array}
$$
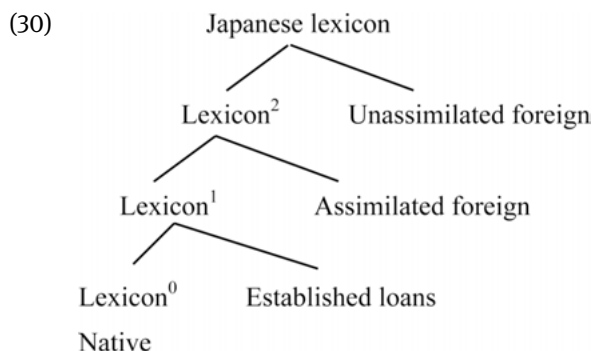
In this conception of the lexicon, different strata are conceived of as forming a gradual and hierarchical stratification characterized by "different degrees of nativization", which are formally represented as a "hierarchy of foreignness, with exceptions to one rule always being exceptions to another rule, but not vice versa" (Itô and Mester 1999: 64). Thus, a particular phonological rule or constraint specifically motivated by the native stratum fails to apply to the three other strata shown in the positions higher in the hierarchy; likewise, a particular phonological rule or constraint that is sensitive to native words and established loans does not apply to the two other classes of assimilated and unassimilated foreign words, for example. An important feature of this model defined by the set inclusion relations of constraint applicability is that the individual constraint domains are primary whereas the stratal structure that emerges from them is only a secondary generalization (Itô and Mester 1999: 70). In other words, the existence of different lexical strata is deemed an epi-phenomenon that follows from the more-or-less systematic clustering of domains that are defined over the applicability of particular rules or constraints.

Attractive as it is, Itô and Mester's core/periphery model could not be upheld as such if we attempt to incorporate into it the morphological properties and word formation processes stated in this chapter as constraint domains. First, as seen in Section 2, only *wago* have the full set of lexical and functional categories; the other strata have missing categories. Next, word formation processes like compounding, derivation, conversion, and reduplication differ in whether they apply or not depending on the strata. Table 8 shows whether or not each lexical stratum includes each type of lexical information or not.

Table 8: Variant applicability of lexical rules

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11
native	+	+	+	+	+	−	+	+	+	+	+
S-J	+	+	+	−	−	+	+	+	+	+	+
foreign	+	+	+	−	−	−	+	+	+	limited	−
mimetic	limited	+	+	−	limited	−	−	+	+	−	+

R1 = Noun, R2 = VN, R3 = AN, R4 = inflected V, R5 = A, R6 = infinite iteration of compounding, R7 = compound head, R8 = compound modifier, R9 = prefix, R10 = suffix, R11 = reduplication

Taking whether or not a word formation rule is applicable as a parameter and marking each parameter as "+" or "−", it is clear that Table 8 does not show a tidy distribution with the unidirectional changes that would be expected from the set inclusion relations of (30). Native words have most properties, but lack at least one – R6. Lexical items considered Sino-Japanese words share many properties with native words, but there are also properties in which they do not participate. What have been considered four independent classes in the past do not each form a single coherent class, but can instead be regarded as forming a gradient along which a word may have a higher or lower affiliation to a given stratum, depending on the greater or lesser number of properties with which it is provided. That is, the four strata do not each constitute a monolithic block; instead, lexical items that have traditionally been considered to belong to the same vocabulary stratum change their membership in the given stratum depending on whether or not each parameter is applicable. Without regard to their historical origin, lexical items in modern Japanese that share a number of lexical characteristics form a single class with the characteristics they share. For example, as seen in the variation in the affixation of the beautific prefix *o-/go-*, a single lexical item, e.g. *tya* 'tea', may have multiple membership, being Sino-Japanese by some standards and native by some other standards.

Thus, when determining vocabulary stratum membership, a lexical item has certain properties, such as the presence or absence of various categories, whether or not certain word formation rules are applicable, or whether or not some lexical phonological constraint applies as criteria and depending on which of many given

standards apply, a degree of membership – a degree of "nativeness" or a degree of "Sino-Japaneseness" – is established incrementally. The greater the degree to which standards characteristic of native words apply, the greater the degree of "nativeness"; the greater the degree to which standards characteristic of mimetics apply, the greater the degree of "mimeticness". Consulting Table 8, except for infinite iteration of compounding, native words have all the properties listed. In this sense, the differences among vocabulary strata can be captured as a difference in the degree of nativization to Japanese, as Itô and Mester (1999) claimed. One important topic for future research may be to determine the properties of vocabulary strata from a synchronic viewpoint through finding the links between various word formation processes and lexical information.

Acknowledgments

References

Akita, Kimi and Natsuko Tsujimura. this volume. Chapter 4: Mimetics. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Halle, Morris. 1973. Prolegomena to a theory of word-formation. *Linguistic Inquiry* 4. 451–464.

Hamano, Shoko. 1998. *The sound symbolic system of Japanese*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Hayashi, Ōki (ed.). 1982. *Zukai nihongo* [Illustrated Japanese]. Tokyo: Kadokawa Shoten.

Irwin, Mark. 2011. *Loanwords in Japanese*. Amsterdam: John Benjamins.

Irwin, Mark. this volume. Chapter 5: English loanwords. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Ishino, Hiroshi. 1992. Gairaigo no zōgoryoku [Productivity of foreign words]. *Nihongogaku* 11(5). 42–49. Tokyo: Meiji Shoin.

Itô, Junko and Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The hand book of Japanese linguistics*, 62–100. Oxford: Blackwell.

Izumi, Kunihisa. 1976. Giseigo, gitaigo no tokushitsu [Characteristics of onomatopoeic words]. In Takao Suzuki (ed.), *Nihongo kōza 4: Nihongo no goi to hyōgen* [A course in Japanese 4: Japanese vocabulary and expressions], 105–151. Tokyo: Taishukan Shoten.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei [Grammar and word formation]*. Tokyo: Hituzi Syobo.

Kageyama, Taro. 2001a. Polymorphism and boundedness in event/entity nominalizations. *Journal of Japanese Linguistics* 17. 29–57.

Kageyema, Taro. 2001b. Word plus: The intersection of words and phrases. In Jeroen van de Weijer and Tetsuo Nishihara (eds.) *Issues in Japanese phonology and morphology*, 245–276. Berlin: Mouton de Gruyter.

Kageyama, Taro. 2002. *Kejime no nai nihongo* [Japanese as an unbounded language]. Tokyo: Iwanami Shoten.

Kageyama, Taro. 2007. Explorations in the conceptual semantics of mimetic verbs. In Bjarke Frellesvig, Masayoshi Shibatani and John C. Smith (eds.), *Current issues in the history and structures of Japanese*, 27–82. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kageyama, Taro. 2010. Variation between endocentric and exocentric word structures. *Lingua* 120. 2405–2423.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kobayashi, Hideki, Kiyo Yamashita and Taro Kageyama. this volume. Chapter 3: Sino-Japanese words. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kokuritsu Kokugo Kenkyūjo [NINJAL]. 1961. *Kokuritsu Kokugo Kenkyūjo hōkoku 20: Dōongo no kenkyū* [NINJAL research report 20: A study of homophonous words]. Downloadable at http://www.ninjal.ac.jp/s data/drep/report nijla/R0020.PDF

Kokuritsu Kokugo Kenkyūjo [NINJAL]. 1964. *Kokuritsu Kokugo Kenkyūjo hōkoku 25: Gendai zasshi 90 shu no yōgo yōji* [NINJAL research report 25: Words and characters used in 90 modern magazines]. Downloadable at http://www.ninjal.ac.jp/s data/drep/report nijla/R0025.PDF

Kokuritsu Kokugo Kenkyujo [NINJAL]. 2006. *Gendai zasshi nihyakumanzi gengochōsa goihyō* [List of vocabulary items in the two-million character database of modern magazines]. Downloadable at http://www.ninjal.ac.jp/archives/goityosa/

Kubozono, Haruo (ed.). 2015. *Handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Labrune, Laurence. 2012. *The phonology of Japanese*. Oxford: Oxford University Press.

Lieber, Rochelle. 1992. Compounding in English. *Rivista di Linguistica* 4(1). 79–96.

Lyons, John. 1977. *Semantics*. 2 volumes. Cambridge: Cambridge University Press.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Miyajima, Tatsuo. 1977. Kindai nihongo ni okeru kango no ichi [The position of Kango in modern Japanese]. In Yasuyuki Suzuki (ed.), *Kokugo kokuji mondai no riron* [Theories about Japanese language and writings], 135–186. Tokyo: Mugi Shobō.

Miyajima, Tatsuo. 1980. Imi-bun'ya to goshu [Semantic fields and vocabulary strata]. *Kokuritsu Kokugo Kenkyūjo hōkoku* [NINJAL research report] 65, 1–16. Downloadable at http://www.ninjal.ac.jp/s data/drep/report nijla/R0065.PDF

Miyake, Tomohiro. 2010. Itizi kango "*suru*"-gata dōshi o megutte [*Suru* verbs with one-character Sino-Japanese words]. In Hiroko Oshima, Akiko Nakajima and Raoul Blin (eds.), *Kango no gengogaku* [The linguistics of Sino-Japanese words], 107–119. Tokyo: Kurosio Publishers.

Miyamoto, Tadao and Hideki Kishimoto. this volume. Light verb constructions with verbal nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word forma tion*. Berlin: De Gruyter Mouton.

Morioka, Kenji. 1970. Wago, kango, gairaigo [Native, Sino-Japanese, and foreign words]. In Kenji Morioka, Masaru Nagano and Yutaka Miyaji (eds.), *Kōza tadashii nihongo 4: Goi* [Correct Japanese Series 4: Vocabulary], 28–60. Tokyo: Meiji Shoin.

Nagano, Akiko and Masaharu Shimada. 2014. Morphological theory and orthography: *Kanji* as a representation of lexemes. *Journal of Linguistics* 50. 323–364.

Nakano, Hiroshi. 1973. Gendai nihongo no onso renzoku no jittai [Sequences of phonemes in present-day Japanese]. *Kokuritsu Kokugo Kenkyūjo hōkoku 49: Denshi keisanki niyoru kokugo kenkyū V* [NINJAL research report 49: Computational study of Japanese V], 94–120. Downloadable at http://www.ninjal.ac.jp/s data/drep/report nijla/R0049.PDF

Namiki, Takayasu and Taro Kageyama. this volume. Chapter 6: Word structure and headedness. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word forma tion*. Berlin: De Gruyter Mouton.

Nihon kokugo daijiten [The grand dictionary of the Japanese language]. 2000–2002. 2nd edition. Tokyo: Shōgakukan.

Nishio, Toraya. 1988. *Gendai goi no kenkyū* [A study of contemporary vocabulary]. Tokyo: Meiji Shoin.

Nishio, Toraya. 2002. Goshu [Vocabulary strata]. In Michiaki Saito (ed.), *Asakura nihongo kōza 4: goi, imi* [Asakura Japanese language series 4: Vocabulary and meaning], 79–109. Tokyo: Asakura Shoten.

Nishiyama, Kunio. 1998. *The morphosyntax and morphophonology of Japanese predicates*. Ithaca, NY: Cornell University dissertation.

Nomura, Masaaki. 1978. Setsujisei jion goki no seikaku [Characteristics of Sino-Japanese suffix-like elements], *Kokuritsu Kokugo Kenkyūjo hōkoku 61: Denshi keisanki niyoru kokugo kenkyū IX* [NINJAL Research report 61: Computational study of Japanese V], 102–138. Downloadable at http://www.ninjal.ac.jp/s data/drep/report nijla/R0061.PDF

Nomura, Masaaki. 1984. Goshu to zōgoryoku [Vocabulary strata and productivity]. *Nihongogaku* 3(9). 40–54. Tokyo: Meiji Shoin.

Saito, Michiaki. 1992a. *Gendai nihongo no gokōseiron teki kenkyū: Go ni okeru katachi to imi* [A Study of the word composition in contemporary Japanese: Form and meaning of words. Tokyo: Hituzi Syobo.

Saito, Michiaki. 1992b. Wago no zōgoryoku [Productivity of native words]. *Nihongogaku* 11(5). 25– 34. Tokyo: Meiji Shoin.

Saito, Michiaki. 2004. *Goiron teki gokōsei ron* [Lexicological approach to word composition]. Tokyo: Hituzi Syobo.

Saito, Michiaki and Masahiko Ishii. (eds.). 2011. *Kore kara no goiron* [Lexicology in prospect]. Tokyo: Hituzi Syobo.

Saito, Michiaki and Masahiko Ishii. 2011. Goiron gaisetsu [Outline of lexicology]. In Michiaki Saito and Masahiko Ishii (eds.), *Kore kara no goiron* [Lexicology in prospect], 11–54. Tokyo: Hituzi Syobo.

Sakakura, Atsuyoshi. 1966. *Gokōsei no kenkyū* [Studies in Word Composition]. Tokyo: Kadokawa Shoten.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Shibatani, Masayoshi, and Taro Kageyama. 1988. Word formation in a modular theory of grammar. *Language* 64. 451–484.

Sibata, Takesi. 1978. *Hōgen no sekai: Kotoba no umareru tokoro* [The world of dialects: Where words are born]. Tokyo: Heibonsha.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Sugioka, Yoko and Takane Ito. this volume. Chapter 10: Derivational affixation in the lexicon and syntax. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Takezawa, Koichi. this volume. Chapter 13: Inflection. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Tamamura, Fumio. 1975. Wago wa zōgoryoku ga yowai ka [Is the productivity of native words low]. In Etsutaro Iwabuchi and Toraya Nishio (eds.), *Gendai nihongo no tango to moji* [Words and orthography of contemporary Japanese], 121–146. Kyoto: Chōbunsha.

Tamamura, Fumio. 1984. *Nihongokyōiku shidō sankōsho 12: Goi no kenkyū to kyōiku* [Guidebook for Japanese-language teaching 12: Teaching of vocabulary I]. Tokyo: Ōkurashō Insatsukyoku.

Tamori, Ikuhiro and Laurence Schourup. 1999. *Onomatope: Keitai to imi* [Onomatopoeic words: Form and meaning]. Tokyo: Kurosio Publishers.

Tateishi, Koichi. 1989. Theoretical implications of Japanese musicians' language. In Jane Fee and Kathryn Hunt (eds.), *Proceedings of the eighth West Coast Conference on Formal Linguistics*, 384–398. Stanford: CSLI Publications.

Teramura, Hideo. 1982. *Nihongo no sintakkusu to imi I* [Syntax and semantics of Japanese I]. Tokyo: Kurosio Publishers.

Tsujimura, Natsuko and Stuart Davis. 2011. A construction approach to innovative verbs in Japanese. *Cognitive Linguistics* 22. 797–823.

Yumoto, Yoko. this volume. Chapter 9: Conversion and deverbal compound nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Hideki Kishimoto and Satoshi Uehara

2 Lexical categories

1 Introduction

Lexical items play different grammatical roles depending on the categories they belong to. Along with the distinction of the vocabulary strata – Native Japanese (NS), Sino-Japanese (SJ), Foreign (F), and Mimetic (M) – the proper classification of lexical categories (word classes) is essential to discovering any grammatical rules and principles in Japanese grammar. In the domestic grammatical tradition, consideration of word classes has a long history starting as early as in the late 1600s, and TŌJŌ Gimon (東条義門, 1786–1843) is generally credited with establishing the fundamental bipartition of word classes into *yōgen* (inflecting class) and *taigen* (non-inflecting class) that persists today. This bipartition is based on the morphological criterion of the presence or absence of inflectional endings, with *yōgen* (inflecting class) comprising verbs and adjectives, and *taigen* (non-inflecting class) nouns. Such a polar opposition, however, will be called into question if there are categories that fall in between the two poles. In fact, contemporary Japanese offers at least two categories that appear to be located between the two, namely Verbal Nouns (VN) and Adjectival Nouns (AN). The present chapter will survey the characteristics of Japanese lexical categories from generative and cognitive perspectives, with due attention to the proper treatment of VNs and ANs in the system of Japanese word classes.

School grammar (Japanese grammar taught at high schools) represents perhaps the most widely spread classification of Japanese lexical categories, which are divided into two major groups: *jiritsugo* or "independent categories" in Table 1 and *fuzokugo* or "dependent categories" in Table 2.

Table 1: Independent-form categories in school grammar

verb (*dōshi*)	adjective (*keiyōshi*)	adjectival noun (*keiyōdōshi*)	noun (*meishi*)
yomu 'read'	*utukusii* 'beautiful'	*sizuka da* 'quiet'	*hito* 'person'
agaru 'climb'	*kanasii* 'sad'	*kirei da* 'pretty'	*anata* 'you'
kuru 'come'	*yoi* 'good'	*derikeeto da* 'delicate'	*terebi* 'TV'
suru 'do'		*barabara da* 'separate'	*wanwan* 'doggie'
adverb (*fukushi*)	prenominal modifier (*rentaishi*)	conjunction (*setsuzokushi*)	interjection (*kantōshi*)
hakkiri 'clearly'	*arayuru* 'every'	*sikasi* 'but'	*hai* 'yes'
sukosi 'a little'	*aru* 'a certain',	*mata* 'or'	*aa* 'oh'
	ano 'that'		
	ookina 'big'		

Table 2: Dependent-form categories in school grammar

auxiliary (*jodōshi*)	particle (*joshi*)
(s)ase 'causative', -(r)are 'passive', -nai/ nu 'negative', -(y)oo 'hortative', -tai 'desiderative', -masu 'polite', -ta/ da 'past, perfect', soo da 'reportive', -mai 'negative expectative', yoo da 'modality (uncertainty)', -rasii 'modality (hearsay)', -da/ desu 'copula', etc.	1. case: *ga* (NOM), *no* (GEN), *o* (ACC), *to* (COM), *kara* (ABL)', *de* (INST), etc. 2. conjunctive: *keredo* 'though', *ga* 'but', *ba* 'if', *node* 'since', etc. 3. adverbial: *wa* (TOP), *mo* 'also', *sae* 'even', *sika* 'only', etc. 4. sentence-final: *na(a)* 'exclamatory', *zo* 'emphatic', *yo* 'assertive', etc.

In regard to vocabulary strata, verbs and adjectives are confined to Native Japanese words. Nouns could be Native Japanese (*kuruma* 'car'), Sino-Japanese (*gakkoo* 'school'), foreign (*terebi* 'TV'), or mimetic (*wanwan* 'doggie') (and verbal nouns, which are not included in the list of lexical categories in school grammar, come from all the lexical strata, as in *dokusyo(-suru)* 'reading (SJ)', *otetudai(-suru)* 'help (NJ)', *tesuto(-suru)* 'test (F)', *tekuteku(-suru)* 'walk (M)'). Adjectival nouns are also found in all the lexical strata, as in *sizuka da/na* 'quiet (NJ)', *ganko da/na* 'stubborn (SJ)', *derikeeto da/na* 'delicate (F)', *barabara da/na* 'separate (M)'.

The classifications in Tables 1 and 2 are far from definitive, and many different views are available in the literature; e.g. in modern Japanese linguistics (couched in the Western linguistic paradigms), verbal nouns are often seen as constituting a major lexical category, alongside verbs, nouns, adjectives, adjectival nouns (see Martin 1975; Shibatani 1990), but they are included in the category of nouns in traditional Japanese grammar. It is not an easy task to sort out Japanese lexical categories, because, in Japanese, inflection/conjugation is exhibited in limited categories, and in many cases, no clear boundaries can be drawn between free and bound elements. Accordingly, a number of issues arise with regard to the question of how lexical categories are distinguished and how many categories should be recognized.

This chapter reviews a number of theoretical and descriptive issues on lexical categories. Section 2 briefly surveys the classification and terms of lexical categories discussed in traditional grammar. Section 3 provides a discussion from the generative perspective, addressing the question of how "adjectival nouns" and "verbal nouns" – neither of which fit into the classical cross-categorial feature system positing only two features – can be defined. It is also shown that certain adjectivally-inflecting elements traditionally classified as dependent auxiliaries count as lexical adjectives, despite their morphological status. Section 4 examines the nature of Japanese lexical categories from the cognitive-typological perspective. The structural organization of the overall lexical categorization of the language and its major lexical categories (Noun, Verb, Adjective, Verbal Noun, and Adjectival Noun) are characterized in terms of cross-linguistic markedness patterns and their functional motivations are identified.

2 Parts of speech in traditional Japanese grammar

This section provides a concise survey of how parts of speech were analyzed by some pioneers of Japanese grammar in the domestic tradition of *Kokugogaku* philology. Parts of speech were mentioned in grammar books written by European missionaries who visited Japan (e.g. Rodriguez 1604–08), and in the pre-modern era (17th to 19th centuries), *Kokugaku* scholars (FUJITANI Nariakira 富士谷成章, MOTOORI Haruniwa 本居春庭, and others) made various attempts to classify inflecting words. Among them, Fujitani (1778) proposed a sophisticated "parts of speech" system of dissecting sentences into four major parts of *na* (noun), *yosohi* (predicate), *ayuhi* (particle, auxiliary), and *kazashi* (adverb and others), while TŌJŌ Gimon (1841), mentioned in Section 1, established the distinction of inflecting and non-inflecting word classes, i.e. *yōgen* and *taigen*. Later, in the late 19th century, ŌTSUKI Fumihiko 大槻文彦 (1889) proposed a system of eight parts of speech in line with the Western tradition of modern linguistics in another important work investigating the notion of "word" and "parts of speech". Subsequently, in the 20th century, the early giants of Japanese grammar (YAMADA Yoshio 山田孝雄, MATSUSHITA Daizaburō 松下大三郎, HASHIMOTO Shinkichi 橋本進吉, and TOKIEDA Motoki 時枝誠記) also proposed different views on parts of speech, which are still influential and debated today.

Among others, Hashimoto's ideas about lexical categories were incorporated in school grammar (Monbushō 1947) and are taught even today. As noted in the Introduction, the inventory of lexical categories includes *dōshi* (verb), *keiyōshi* (adjective), *keiyōdōshi* (adjectival noun), *meishi* (noun), *fukushi* (adverb), *rentaishi* (prenominal modifier), *setsuzokushi* (conjunction), *kantōshi* (interjection), *jodōshi* (auxiliary) and *joshi* (particle). The listed categories are categorized by two distinct sets of criteria, morphological and distributional (or syntactic), as represented in the taxonomy structure of (1), where M stands for morphological criteria and D for distributional criteria.[1]

---

1 Categories are distinguished by making use of a feature system here, although this is not a common practice in school grammar. In generative linguistic research, and also in traditional Japanese grammar, lexical/grammatical categories (word classes, parts of speech) are distinguished on the basis of their shared morphological and syntactic properties. Semantic criteria are not commonly used for this purpose because, in many, if not all, cases, they are not considered to demarcate categories in a reliable manner. For instance, the semantic criterion taking noun and verb to denote an entity and an event, respectively, does not distinguish *nagare* 'flow (n.)' as in *mizu no nagare* 'water flow' from *nagareru* 'flow (v.)', both of which denote events in one way or another; *nagare* refers to an event, and yet functions as a noun.

(1)

```
                                    Word
                       _____/    _____
                      /                              \
              [−dependent] (M)                  [+dependent] (M)
              ____/      \____                     ___/    \___
             /                \                   /            \
     [+inflected] (M)   [−inflected] (M)   [+inflected]   [−inflected]
            |            ____/     \____          |             |
     [+predicate] (D)  [+subject] (D)  [−subject] (D)           |
            |             |          ___/    \___               |
         (yōgen)       (taigen) [+modifier] (D) [−modifier] (D) |
            |             |          |             |            |
          dōshi         meishi    fukushi     setsuzokushi   jodōshi   joshi
         keiyōshi                 rentaishi    kantōshi
        keiyōdōshi
```
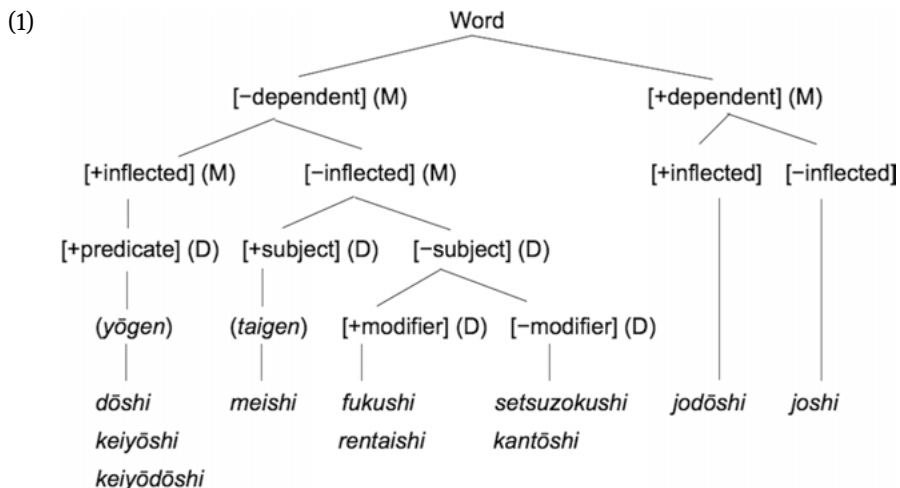
In (1), words are first divided into two major groups, depending on whether they are dependent or independent ([±dependent]). Both categories are further partitioned into two classes, according to whether they inflect or not ([±inflected]). All the inflecting categories are identified as predicates [+predicate], and non-inflecting categories are divided into two sub-types, depending on whether or not they can function as subjects ([±subject]). Non-inflecting categories that cannot act as subjects are further divided into two classes, depending on whether they can serve as modifiers [±modifier].

The predicative group marked with [+predicate], which comprises verbs, adjectives, and adjectival nouns, is traditionally labeled *yōgen* 'predicative', and nouns are the sole member of the class labeled *taigen* 'nominal'. *Yōgen* is divided into two types, as in (2a), which are distinguished according to whether or not they have imperative inflections (represented as [±imperative]).

(2) a.
```
                yōgen
            ___/    \___
           /            \
   [+imperative] (D) [−imperative] (D)
         |                  |
       dōshi            keiyōshi
                        keiyōdōshi
```
b.
```
                [+modifier]
             ___/       \___
            /               \
   [+adverbial] (D)   [+prenominal] (D)
         |                  |
      fukushi           rentaishi
```
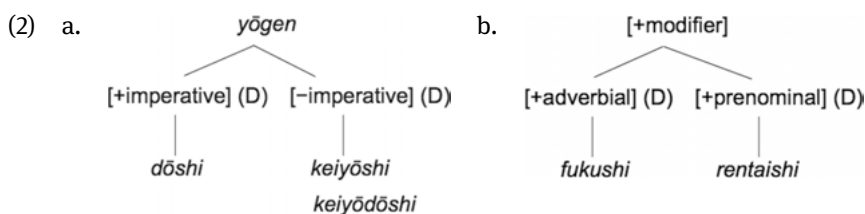
Modifiers are divided into *fukushi* (adverbial) and *rentaishi* (prenominal modifier), which are distinguished by the kind of element modified. *Setsuzokushi* is used for connecting clauses, and *kantōshi* for forming 'exclamatives', which may or may not be single-word sentences.

Of the categories listed above, *keiyōdōshi* (adjectival noun), *rentaishi* (prenominal modifier), *jodōshi* (auxiliary), and *joshi* (particle) call for special mention, as they are not commonly found in European languages or are different from their European counterparts. *Keiyōdōshi*, referred to in this chapter as "adjectival noun" (a term coined by Martin 1975), is chosen for the sake of consistency over competing English translations, and the expression consists of a stem and an inflectional ending. The inflectional endings of *keiyōdōshi* (e.g. adnominal *na* and conclusive *da*) are morphologically distinct from those of adjectives (e.g. adnominal and conclusive -*i*). As explained below, the category of *keiyōdōshi* poses a perennial question for both *Kokugogaku* and Japanese linguistics in line with the Western tradition. *Rentaishi* ('prenominal modifier'), which is often given the alternative term 'prenominal adjective', include not only demonstratives like *ano* 'that', but also such adnominal expressions as *aru* 'a certain', *iwayuru* 'so-called', *taisita* 'considerable', *tiisana* 'tiny', and *ookina* 'big' that occur only in prenominal position. *Jodōshi* ('auxiliary'), perhaps the most undeveloped, problematic category in traditional grammar, is actually a conglomeration of heterogeneous dependent elements that are agglutinated to the preceding predicative elements, such as suffixes of passive, causative, honorification, spontaneity, and potential, as well as functional endings of tense, negation, modality (see Chapter 13 [Takezawa, this volume] for inflection). Finally, various particles labeled as *joshi* are dependent elements that follow noun phrases, clauses, and the like (but are not agglutinated to them).

Although the classifications in school grammar are said to be largely based on work by Hashimoto (1934, 1948), Hashimoto (1934) actually presented a finer-grained classification including not only *jiritsugo* (independent category) but also *fuzokugo* (dependent category), as in *yōgen* (verb, adjective), *taigen* (noun, pronoun, numeral), *fukushi* (adverb), *rentaishi* (prenominal modifier), *setsuzokushi* (conjunction), *kantōshi* (interjection), *jodōshi* (auxiliary), as well as *joshi* (particle), a category that includes various types of particles such as adverbial, nominal, conjunctive, coordinate, quasi-adverb, case, *kakari*, sentence-final, and interjective particles. (In school grammar, *keiyōdōshi* is treated as an established lexical category (cf. Hashimoto 1948), but in his earlier work (Hashimoto 1934), Hashimoto eschews including *keiyōdōshi* in the list of *yōgen* as an independent lexical category in modern Japanese, reserving the judgment on the issue as to whether it should be analyzed as a simple word or a complex expression, although he argues explicitly for its unique status there.)

The question of how words are classified involves a fundamental and yet highly subtle issue, one that remains far from settled even today. The debates in the *Kokugogaku* tradition largely center around the two somewhat related questions: (i) whether it is plausible to posit *keiyōdōshi* (adjectival noun) as an independent category, and (ii) how many parts of speech need to be recognized. Answers to these questions may vary, depending not only on how much emphasis one places on what kinds of criteria (morphological or distributional), but also on what kind of language one analyzes (classical or modern). In traditional grammar, it was customary to

consider category classifications based on classical Japanese or a mixture of classical and modern Japanese. The first question concerning *keiyōdōshi* is rooted in the structure of classical Japanese. The name *keiyōdōshi* (lit. 'adjectival verb') derives from its characteristic forms in classical Japanese, X-*nari* and X-*tari*, which exhibited verbal inflections and inflected even for the imperative. In modern Japanese, where the classical *nari* and *tari* have disappeared, however, *keiyōdōshi* has lost such erstwhile verbal behavior. Hashimoto (1948) claimed that *keiyōdōshi* constitutes a lexical category standing all on its own (although this view was only implied in Hashimoto (1934)). On the other hand, Yamada (1908, 1936) considered this categorization unnecessary on the grounds that the classical *nari* and *tari* originate historically from *ni ari* [LOC be] and *to ari* [COMP be], where *ari* 'be' is a verb of existence. On Yamada's view, *keiyōdōshi* is thus composed of what he calls *fukuyōgen* 'sub-predicate' (words that are neither inflected as predicates nor function as grammatical arguments) followed by predicates of existence (*sonzaishi* 'existential' in Yamada's terminology).

In modern Japanese, *keiyōdōshi* is "inflected" with the endings *na* in adnominal form and *da* in conclusive form. The ending *na* comes from the classical combined form *ni* + *ari* [LOC + be], and the *da* ending is identical in form to the copula *da* as in *Nihon no syuto wa Tookyoo da* [Japan GEN capital TOP Tokyo COP] 'The capital of Japan is Tokyo.' This and other considerations led Tokieda (1950) to analyze *keiyōdōshi* into a special type of nominal followed by the copula. The view that the inflectional ending of *keiyōdōshi* is nothing but the copula is shared by Bloch (1946) and Okutsu (1978). On the other hand, Mikami (1953), inspired by Sakuma (1951), holds that adjectives and adjectival nouns make up a larger class of adjectivals, the former represented as i-*keiyōshi* (*i*-adjective) and the latter as na-*keiyōshi* (*na*-adjective). The terms "i-*keiyōshi*" and "na-*keiyōshi*" are standardly used today in textbooks for teaching Japanese as a foreign language, whereas "*keiyōshi* (adjective)" and "*keiyōdōshi* (adjectival noun)" are taught in domestic school grammar. The nature of the two adjectival categories will be discussed in great detail in the subsequent sections of this chapter (see also Chapter 14 [Kageyama, this volume] for the morphological status of the *na/da* ending in *keiyōdōshi*).

The writings of such great grammarians as Hashimoto, Yamada, Matsushita, and Tokieda present diverse answers to the second question of how many parts of speech should be distinguished. Although there is general agreement on the usefulness of the fundamental distinction between *taigen* (nominal) and *yōgen* (predicative), these grammarians disagree on the classification of major categories and functional categories. Yamada (1908, 1936) identifies four categories consisting of *taigen*, *yōgen*, *fukushi* (adverb), and *joshi* (particle) but does not regard auxiliaries as an independent class. Matsushita (1924, 1930) does not identify auxiliaries and particles as categories, because they are dependent, thus upholding the classification of *meishi* (noun), *dōshi* (verb), *fukutaishi* (adjective), *fukushi* (adverb), and *kandōshi* (exclama-

tive). Tokieda (1950) includes *taigen*, *yōgen*, *daimeishi* (pronoun), *rentaishi* (prenominal modifier), and *fukushi* (adverb) as "core" lexical categories.

The various classifications that have been proposed in *Kokugogaku* are not without problems from the perspective of modern linguistics. One characteristic feature of traditional Japanese grammar is that lexical categories are first divided into two major classes according to whether they count as free or bound forms in morphological terms. In fact, in all the classifications noted above, morphological criteria take precedence over distributional criteria, but it is not entirely clear whether this should be considered the optimal way of characterizing lexical categories.

In particular, the agglutinative character of Japanese makes it difficult to tell whether dependent elements appearing after main predicates should be analyzed as auxiliaries or inflectional suffixes. A number of issues arise from this fact. One such issue concerns the question of whether the element *da* combined with adjectival noun stems should be regarded as a copula or an inflectional suffix (see Section 4). Copulas are dependent elements, but often regarded as non-affixal, unlike other auxiliaries, leading to the issue over whether adjectival nouns constitute single or complex words. Another question is how the past (or perfective) tense marker should be classified. Traditionally, it is often classified as an auxiliary, distinct from the present tense element, which is taken to be part of an inflectional form, while Bloch (1946) identifies both present and past tense markers as inflectional suffixes. The typological character of Japanese as an agglutinative language also raises an issue over whether there should be a tight correlation between the distinction of "agglutinated" versus "non-agglutinated (free)" forms and the "lexical" versus "functional" distinction (see Section 3).

3 Lexical categories in the generative perspective

We are now in a position to discuss how "adjectival nouns" and "verbal nouns" can be described from the generative perspective. It is shown in this section that adjectivally-inflecting auxiliaries sometimes display syntactic behavior shared with lexical adjectives, and that in Japanese, the categorical distinction of "lexical" and "functional" categories is not tightly correlated with the morphological status of words.

3.1 Criteria for classifying major categories

In generative grammar, mainly two kinds of distributional criteria – morphological and syntactic – are used to identify categories (see, e.g. Carnie 2007). The morphological criteria are based on inflections and derivations. Different inflectional forms

are usually found depending on their categories: in Japanese, verbs have inflectional forms distinct from those found in adjectives or adjectival nouns, and nouns do not inflect for person and number. The derivational criteria work more or less in a similar way. In Japanese, adjectives as well as adjectival nouns can be identified as falling into the adjective class by looking at whether the suffix −*sa* can be attached, since it derives a noun from an adjectival expression. The syntactic (or distributional) criteria are also usable for categorizing words, because different categories have different syntactic distributions. A noun can appear in the frame   *ga nai* [NOM NEG] 'there is no   ', but the same slot cannot be filled by an adjective, an adjectival noun, a verb, or a postposition; e.g., *hon* 'book' is identified as a noun, because it can appear in this frame.

Lexical categories consist of a number of word classes, which include the major categories of noun (N), verb (V), adjective (A), preposition/postposition (P) (i.e. content words), as well as several other minor categories (i.e. function words). In Chomsky (1970), the lexical categories are not conceived of as primitives, but are defined by implementing cross-categorial features, i.e. the four major categories of V, N, A, and P in English are characterized in terms of the features [±V, ±N].

(3) a. V: [+V, −N]   b. N: [−V, +N]   c. A: [+V, +N]   d. P: [−V, −N]

One obvious advantage of the cross-categorial feature system, compared with positing just distinct categories (or having simple binary classifications), is that it allows us to capture commonalities across categories. In Chomsky's system, the maximal number of categories that can be defined with the features is four. On the empirical level, however, it is obvious that the cross-categorial feature system used to define the categories in English often cannot be straightforwardly carried over to other languages, including Japanese. (Even in English, no consensus has been reached about the classification of major lexical categories (content words), and in particular, the exact categorical status of 'preposition' is often called into question; Emonds (1985) and Baker (2003) suggest that prepositions fall into the class of functional categories).

Japanese is often seen as possessing verbal nouns and adjectival nouns as major lexical categories, alongside verbs, nouns, and adjectives (see Martin 1975, Shibatani 1990). Apparently, Japanese has more categories than can be defined by the features, so the feature system has raised an issue over how lexical categories should be defined. Kageyama (1982) suggests that the lexical categories in Japanese be defined with reference to the feature [±A], in addition to [±V, ±N] (see also Kageyama 1993).

(4) a. verb: [+V, −N, −A]   b. noun: [−V, +N, −A]
    c. adjective: [−V, −N, +A]   d. adjectival noun: [−V, +N, +A]
    e. verbal noun: [+V, +N, −A]

Note that in (4), adjectives are characterized as [−N, −V], unlike Chomsky's feature characterization [+V, +N]. They are characterized negatively by the respective features because [+A] is added. In Kageyama (1982), adjectival nouns are claimed to constitute an independent lexical category, on the grounds that they display dual behavior as nouns and adjectives. To be concrete, adjectival nouns pattern with adjectives, in undergoing nominalization with the addition of the suffix −*sa*, which derives nouns from adjectives.

(5) a. *sizuka-sa*   b. *utukusi-sa*   c. **tabe-sa*   d. **enpitu-sa*
         quiet-NMLZ   beautiful-NMLZ   eat-NMLZ   pencil-NMLZ
         'quietness'   'beauty'   'eat-ness'   'pencil-ness'

On the other hand, adjectival nouns differ from adjectives, in that the former, but not the latter, allow their stems to be combine with the affix −*rasii*. Adjectival nouns pattern with nouns, since both of their stems can be combined with *-rasii*, as in (6).

(6) a. *sizuka-rasii*   b. **utukusi-rasii*   c. *gakkoo-rasii*
         quiet-looking   beautiful-looking   school-looking
         'look quiet'   'look beautiful'   'look like a school'

Neither nouns and adjectives show exactly the same behavior as adjectival nouns, so adjectival nouns are characterized as [−V, +N, +A] in Kageyama's system.

In Kageyama (1982), verbal nouns are regarded as constituting a category distinct from the categories of noun and adjective, on the basis that verbal nouns serve not only as arguments occurring with case particles, but also as predicates when combined with *suru* 'do' (and the latter is not possible with nouns; e.g. **enpitu-suru* 'pencil-do'). (Note that verbal nouns can behave as predicates without *suru* when they are combined with the suffix −*tyuu*, *-sai*, *-ori* and the like (see Kageyama 1993)).

(7) a. *soodan-suru*   b. *soodan*   *o*   *motikakeru*
         consultation-do   consultation   ACC   bring.up
         'consult'   'ask for consultation'

Accordingly, verbal nouns are assigned the feature [+V, +N, −A], which is distinct from nouns specified as [−V, +N, −A]. In his analysis, the commonalities shared by nouns and verbal nouns are captured by the feature [+N], whereas the properties which verbal nouns share with verbs are characterized by the feature [+V] (and verbal nouns and verbs are assigned [−A], as they do not possess properties shared with adjectives).

While Kageyama (1982) postulates the feature [±A], which is not implemented in Chomsky (1970), Miyagawa (1987) advances an alternative analysis in an attempt to

dispense with this extra adjectival feature and proposes an alternative characterization of the lexical categories.

(8) a. verb: [+V, −N]   b.   noun, verbal noun: [−V, +N]

c.   adjective: [+V]   d.   adjectival noun: [+V, +N]

In Miyagawa's analysis, adjectives are conceived of as possessing the feature [+V], while the feature [±N] being neutralized. In contrast, adjectival nouns are associated with both [+V] and [+N]. One indication that adjectival nouns, but not adjectives, are associated with [+N] is found in the adjectival nouns' ability to occur with copula, in an analogous way with nouns, as in (9).[2]

(9) a. *kirei   da*   b.   *sensei   da*   c. **utukusi   da*
          pretty   COP   teacher   COP   beautiful   COP
          'be pretty'   'be a teacher'   'be beautiful'

Further, on the basis that adjectival nouns can be combined with *soo* 'looking, likely', alongside verbs and adjectives, Miyagawa defends the position that adjectival nouns are assigned the feature [+V].

(10) a. *sizuka-soo*   b. *utukusi-soo*   c. *tabe-soo*   d. **gakkoo-soo*
          quiet-looking   beautiful-looking   eat-looking   school-looking
          'likely quiet'   'likely beautiful'   'likely to eat'   'likely school'

Miyagawa claims that the generalization that adjectival nouns, adjectives, and verbs can combine with *soo* 'looking, likely' can be optimally characterized by assuming that these three categories have the feature [+V] in common.

In Miyagawa's analysis, adjectival nouns are characterized as having [+N, +V], whereas adjectives, which are marked by the feature [+V], are neutral with respect to the feature [±N]. One problem with the proposed characterization is that the nominalizing affix *-sa* attaches to both adjectival nouns and adjectives, to the exclusion of verbs and nouns, as we have seen in (5). Miyagawa claims that this distribution can be captured on the assumption that *-sa* can be attached to a category with [+V] that does not assign any case (i.e. verbs belong to the category that assigns case, but adjectival nouns and adjectives do not). Under this proposal, adjectives

---

2 The adnominal form of the copula *da* is *no* when the preceding element is a noun, but the copula has the form *na* when it is preceded by an adjectival noun. In Miyagawa's analysis, the difference in morphological inflection between nouns and adjectival nouns is taken care of by adjustment rules. Miyagawa provides two more arguments which he claim can be used to distinguish adjectival nouns from adjectives, i.e. the attachment of *mitai* 'seem' and *reba* 'if'. Ohkado (1991) points out some problems in using them as diagnostics, however.

and adjectival nouns can be classed together by virtue of the case requirement imposed on the categories to which *-sa* attaches, without recourse to the additional feature [±A].

Another claim made by Miyagawa is that verbal nouns are equated with nouns in terms of categorical features, the difference being reduced to the question of theta role assignment. According to Miyagawa, the fact that *suru* 'do' directly attaches to verbal nouns, but not nouns like *enpitu* (**enpitu-suru* [pencil-do]), comes from the condition that *suru* needs to inherit a theta role from a combined theta role assigner, which he assumes is imposed independently of the category of the theta role assigner. If a noun does not have any theta role to assign, as in *enpitu*, the resulting N-*suru* is not capable of acting as a predicate, and hence is unacceptable. This analysis allows us to account for the fact that verbal nouns (equipped with theta roles to be assigned), but not simple nouns, can occur with the verb *suru*, without any need to set up the extra lexical category of verbal noun.

Both Kageyama and Miyagawa attempt to capture generalizations across the major lexical categories (content words) in terms of cross-categorial features. Even though the issue concerns the status of major lexical categories, there are also minor categories (function words), which are often identified with the categorical feature [+F] (or [−L]) instead of [+L] assigned to a major lexical category (Abney 1987; Fukui 1986). Needless to say, just as English has minor functional categories such as pronouns, auxiliary verbs, etc., so too does Japanese, which leads to another issue. We will turn to this discussion in the next section.

3.2 From major to minor categories

In Japanese, auxiliaries, which are inflecting suffixal/agglutinative elements attaching to predicates, are dependent categories, and many of them are likely to be categorized as functional categories. Nevertheless, it is also true that some, if not all, have an obvious link to major lexical categories, partly because the former have often been derived from the latter via grammaticalization (see Hopper and Traugott 1993), as can be inferred from the fact that auxiliaries often retain their morphological forms of the original categories from which they are derived; some display adjectival properties (or inflections) (e.g. the desiderative *ta(i)* 'want', the hearsay *rasi(i)* 'likely'), and others display verbal behavior (e.g. the passive *rare(ru)*, the causative *sase(ru)*).

The fact that auxiliaries can be divided into verbal and adjectival types can also be discerned conspicuously by looking at what supportive verb is inserted when tense is separated from them, as in (11).

(11) a. *{yomi/yom-are/yom-ase}*   *mo*   *su-ru*
          {read/read-PASS/read-CAUSE}   also   do-PRS
          (lit.) 'do also {read/be read/cause to read}'

      b. *{utukusiku/nomi-taku}*   *mo*   *ar-u*
          {beautiful/drink-want}   also   be-PRS
          (lit.) 'be also {beautiful/want to drink}'

The supportive verb *suru* 'do' is inserted to the left of tense for morphological support when an adverbial particle separates the tense from the passive *rare* and the causative *sase*, just like the main verb *yomu* 'read'. This fact suggests that the auxiliaries *rare* and *sase* belong to the verbal class (cf. Kishimoto 2013). In contrast, when tense is separated from an adjective like *utukusii* 'beautiful' or the desiderative auxiliary *tai*, the supportive verb *aru* appears to the left of the tense marker. Since both elements pattern together, this fact shows that the auxiliary *tai* belongs to the adjectival type.

     In traditional Japanese linguistics, inflecting elements that are dependent, i.e. those that cannot appear in isolation, are categorized as auxiliaries (by definition). Given that major lexical categories are independent forms, it is tempting to think that auxiliaries are assimilated to functional categories. Nevertheless, the fact of the matter is that auxiliaries constitute a heterogeneous class, and it will be shown below that some agglutinated auxiliaries showing adjectival inflection may be categorized as lexical adjectives. By appealing to distributional/syntactic criteria, it is argued that the morphological distinction of "agglutinated" versus "non-agglutinated (free)" forms does not directly correlate with the distinction between "lexical" and "functional" categories, and that bound morphemes can be lexical (or content) morphemes, which should fall into the major lexical class.

     To be concrete, *tai* 'want', which is categorized as an auxiliary in traditional Japanese linguistics (Konoshima 1973; Kitahara 1981; and many others), is similar in meaning to the adjective *hosii* 'want' (or the verb *hossuru* 'want'), the most obvious difference between the two being that the former is an agglutinative element [+Aggl], and the latter a free form [−Aggl]. The auxiliary *tai* needs to combine with a verb, but *hosii* stands as a syntactically independent element, as in (12).

(12) a. *Watasi*   *wa*   *okane*   *ga*   *hosi-i.*
          1.sg   TOP   money   NOM   want PRS
          'I want money.'

      b. *Watasi*   *wa*   *osake*   *ga*   *nomi-ta-i.*
          1.sg   TOP   sake   NOM   drink-want-PRS
          'I want to drink sake.'

The auxiliary *tai* can be separated from the preceding verb with an appropriate morphological adjustment, i.e. with the addition of dummy verb *suru* to the left of *tai* (e.g. *kangae-tai* [think-want] 'want to think' → *kangae wa si-tai* [think TOP do-want]), and differs from a derivational affix like *-sa* (as in *kata-sa* [solid-ness]), which can never be separated from its host. Both *tai* and *hosii*, despite the difference in their morphological status, function as lexical adjectives, i.e. predicates specified for [+Adj, +Pred], as confirmed by the fact that they can appear as predicates embedded under the verb *omou* 'think', which takes a small clause complement.

(13) a. *Watasi wa [sono okane o amari hosiku] omowa-nakat-ta.*
      1.sg TOP that money ACC much want think-NEG PST
      'I did not want that money very much.'

   b. *Watasi wa [osake o nomi-taku] omo-u.*
      1.sg TOP sake ACC drink-want think-PRS
      'I think that I want to drink sake.'

Note that the small clause complement selected by *omou* can have an adjective (and also an adjectival noun) as its predicate (Kishimoto 2007, 2008).

(14) *Ken wa [sono kodomo o {kawaiku/sinsetu-ni}] omot-ta.*
    Ken TOP that child ACC {cute/kind} think-PST
    'Ken thought that child {cute/kind}.'

Example (13b) is acceptable because *nomi-tai* 'want to drink' is an adjectival predicate, with the structure *complement verb + lexical adjective* (*tai*), i.e. the auxiliary *tai* in (13b) functions as a fully lexical adjective with the feature [+Adj, +Pred, +Aggl], despite the fact that it is a morphologically dependent (agglutinative) element. By contrast, (15) is excluded because the embedded predicate has the sequence of *verb + function word* (*nai*), i.e. the verb is negated by the function word *nai*.

(15) **Ken wa [sono kuruma o ure-naku] omot-ta.*
    Ken TOP that car ACC sell.can-NEG think-PST
    'Ken thought that car unlikely to be sold.'

In (15), the negative *nai*, which inflects like an adjective, is a functional category, which can be labeled as [−Pred, +Aggl] (i.e. [−Pred] → [−Adj, −Pred]), and thus, it is not allowed to occur in the small-clause complement when it is combined with a verb.

The difference in the grammatical status of *tai* and *nai* can also be confirmed by embedding them under *hosii* 'want'. When *nomi-tai* and *noma-nai* are embedded under *hosii*, a difference in acceptability emerges, as in (16).

(16) a. *Watasi* *wa* *[Ken* *ga* *osake* *ga* *nomi-taku-te]* *hosi-i.*
     1.sg   TOP Ken NOM sake   NOM drink-want-GER want PRS
     'I want Ken to want to drink sake.'

    b. *Watasi* *wa* *[Ken* *ga* *osake* *o* *noma-nai-de]* *hosi-i.*
     1.sg   TOP Ken NOM sake   ACC drink-NEG GER want PRS
     'I want Ken not to drink sake.'

Note that when a negated verb is embedded under *hosii*, the *nai-de* form instead of the *naku-te* form must be used. The verbal *te*-form *nai-de* can be formed only when it is preceded by a verb, but the *naku-te* form is available for both negated verbs and adjectives (see e.g. Kuno 1973). The contrast in acceptability between (16a) and (16b) comes from the requirement that the complement clause selected by *hosii* should be verbal, but not adjectival. This is confirmed by (17).

(17) a. *Watasi* *wa* *[sono* *kuruma* *ga* *ure-te]* *hosi-i.*
     1.sg   TOP that   car   NOM sell.can-GER want PRS
     'I want that car to be sold.'

    b. *Watasi* *wa* *[kodomo* *ga* *itumo* *kawaiku-te]* *hosi-i.*
     1.sg   TOP child   NOM always cute-GER   want-PRS
     'I want the child to be cute at all times.'

Thus, the difference in acceptability between (16a) and (16b) suggests that *nomi-tai* has the constituent structure *verb + adjective*, and *noma-nai*, the structure *verb + function word*. The data show that certain adjectivally-inflecting auxiliaries may be construed as lexical adjectives, even though they are morphologically dependent.

    Notably, predicates identified as morphologically free can be either lexical or functional. In traditional grammar, the existential/possessive *nai*, which is the negative form of the verb *aru*, is claimed to be an "adjective" mainly for morphological reasons, but the examples in (18) illustrate that it does not function as a lexical adjective.

(18) a. *Ken* *wa* *[soko* *ni* *hon* *o* *naku]* *omot-ta.*
     Ken TOP there LOC book ACC NEG   think-PST
     'Ken thought the book not to be there'

    b. *Watasi* *wa* *[soko* *ni* *hon* *ga* *{naku-te/nai-de}]* *hosi-i.*
     1.sg   TOP there LOC book NOM {NEG GER/NEG GER} want PRS
     'I want the book to not to be there.'

The existential/possessive *nai* is not allowed to appear in the context where an adjective can appear, nor can it occur in the syntactic context where a verb is

allowed. This suggests that the existential/possessive *nai* should belong to the functional class, despite the fact that it is a non-agglutinated form morphologically; hence this element should be marked as [–Pred, –Aggl].

In the literature on Japanese, a number of opinions are available as to how *nai* should be related to *aru* (see Hashimoto 1969; Yamaguchi 2004; Kato 1985; and many others). But if the existential/possessive *nai* – the negative counterpart of *aru* 'be' – is a functional element, as noted above, a reasonable analysis would be that, as suggested by Kato (1985), the negative form *nai* is derived from *ara-nai* by dropping the verb part *ara-*, owing to the fact that *aru* somehow lacks an inflectional form **ara-* in contemporary Japanese. It is worth noting here that archaic negative expressions like *ara-nu* and *ara-zu* do inclde the verb form *ara-*, which must be deleted when combined with *nai*. Note further that the verbal negator *nai* can only have the *te*-form *naku-te*, unlike the verbal negation which can have both *naku-te* and *nai-de*, because the verb is dropped.

The regular negator *nai* combined with verbs does not have the properties of a lexical adjective, despite its adjectival inflection, and serves as a functional predicate. It is plausible to say here that this type of *nai* acquires status as a functional category via the process of decategorialization – a shift from [+Adj, +Pred] to [–Adj, +Pred] (i.e. the loss of 'adjective' status), and further, to [–Pred] (the loss of 'predicate' status) In this case, the functional shift of *nai* has taken place while retaining its inflection. (Note that major lexical categories are often grammaticalized into minor categories while preserving their inflectional patterns; see e.g. Brinton and Traugott 2005). Interestingly, there are cases where negative *nai* retains the status as a lexical adjective, as exemplified by *warikire-nai* 'unsatisfactory' and *abunage ga nai* 'without danger'.

(19) a. *Ken wa [sono kettei   o   warikire-naku] omot-ta.*
     Ken TOP that   decision ACC satisfy-NEG   think-PST
     'Ken thought that decision to be unsatisfactory.'

    b. **Watasi wa [Ken ni   sono kettei   ga   warikire-nai-de] hosi-i.*
       1.sg   TOP Ken DAT that decision NOM satisfy-NEG GER want PRS
       'I want Ken to want to drink sake.'

    c. *Watasi wa [kare no   unten   o   abunage ga   naku] omo-u.*
       1.sg   TOP he   GEN driving ACC danger NOM null   think-PRS
       'I think his driving without danger.'

As shown in (19a, c), *warikire-nai* and *abunage ga nai* can be legitimately embedded as a small clause predicate under *omou* (and *warikire-nai* cannot be embedded under *hosii*, as in (19b)). The data suggest that negative *nai*s associated with these

expressions function as lexical adjectives, i.e. [+Adj, +Pred]. Nevertheless, their morphological status differs, as we can see from (20).[3]

(20) a. *warikire mo si-nai*   b. *(nan no) abunage mo nai*
      satisfy also do-NEG   any GEN danger also null
      'not also be satisfied'   'also without (any) danger'

In both cases in (20), an adverbial particle can be added to the front of *nai*. In (20a), the supportive verb *suru* occurs to the left of *nai*, which shows that *nai* is an agglutinated element, i.e. [+Aggl]. In (20b), no supportive element is necessary, since *nai* here is a non-agglutinated form, i.e. [–Aggl], just like the existential/possessive *nai*. Both instances of *nai* in (20) are categorized as adjectives by the syntactic criteria, as verified by (19a, c). This shows that *nai* can be a lexical adjective, regardless of whether it is agglutinated or not, i.e. *nai* appearing in *warikire-nai* is specified as [+Adj, +Pred, +Aggl], and *nai* appearing in *abunage ga nai* as [+Adj, +Pred, –Aggl].

    It is worthy of note that Japanese has a deverbal predicate as well. The predicate *iru* 'need' provides a case in point. This predicate shows verbal inflection. Even though *iru* counts as a verbal, at least, in morphological terms (cf. Backhouse 2009), it cannot be embedded under *hosii* 'want'.

(21) **Watasi wa [Ken {ga/ni}   okane ga it-te]   hosi-i.*
      1.sg   TOP Ken {NOM/DAT} money NOM need-GER want PRS
      'I want Ken to need money.'

As discussed by Kishimoto (2005), any kind of lexical verb can appear in the embedded clause introduced by *hosii*.[4] Then, (21) shows that *iru* 'need' does not act as a lexical verb, even though it behaves as a predicate with an argument structure determining the thematic status of their arguments. This suggests that *iru* 'need' serves as a predicate devoid of its categorical property as a verb, i.e. [–V, +Pred,

---

3 These adjectival expressions do not have affirmative forms, as in **warikireru* and **abunage ga aru*. *Abunage ga nai* serves as an idiomatic adjective, so it can be easily embedded under *omou* even though *abunage* bears nominative case marking. On the other hand, when *nai* serves as a grammatical negator, it has an affirmative counterpart, as (ia) shows, and the clause cannot be embedded under *omou* 'think', as in (ib).

(i) a. *Kare ga yuuki ga {na i/ar u}.*
      he NOM courage NOM {NEG-PRS/be PRS}
      'He {does not have/has} courage.'

   b. **Watasi wa [kare o   yuuki ga naku] omo u.*
      1.sg TOP he ACC courage NOM NEG think PRS
      'I think him without courage.'

4 In (21), any type of lexical verb can appear in the subordinate clause, because this clause involves simple embedding. Thus, non self controlable verbs, as well as stative verbs, can be used as the predicates of the embedded clause, as shown in (i).

−Aggl], as opposed to an ordinary verb labeled as [+V, +Pred, −Aggl]. Furthermore, the negative marker *nai* combined with *iru* 'need' behaves differently from the regular negator *nai* (associated with ordinary verbs), which is identified as a functional category, as shown in (22).

(22) a. **Watasi   wa   [Ken   {ga/ni}   okane   ga   ira-nai-de]   hosi-i.*
      1.sg   TOP Ken {NOM/DAT} money NOM need NEG GER want PRS
      'I want Ken to need money.'

    b. *Watasi   wa   [sono okane   o   ira-naku]   omot-ta.*
      1.sg   TOP that   money ACC need-NEG think-PST
      'I thought that money unnecessary.'

Since *ira-nai* 'need not' can be embedded under *omou*, but not under *hosii*, the negative *nai* associated with *iru* 'need' must count as a lexical adjective, i.e. [+A, +Pred, +Aggl], illustrating that *nai* can be a lexical adjective even if it is a bound form. The facts of the negative *nai* illustrate that there is no necessary connection between the "lexical" versus "functional" distinction, on the one hand, and the "agglutinated" versus "non-agglutinated" distinction, on the other.

    Let us now turn to the discussion of the predicate *sugiru*, which carries the meaning of 'exceed' or 'pass', for further illustration of the fact that categories cannot be identified solely in morphological terms. Note first that *sugiru* can be used as a main verb, as in (23).

(23) a. *Zikan   ga   sugi-ru.* b. *Hatugen   no   do   ga   sugi-ru.*
      time   NOM pass-PRS   statement GEN degree NOM excess-PRS
      'Time passes.'   'The statement is too excessive.'

Sugiru can also appear as part of compound verbs, as shown in (24).

(i) a. *Watasi   wa   [gohan   ga   suguni   deki te]   hosi i.*
        1.sg   TOP rice   NOM immediately make GER want-PRS
        'I want the meal to be ready immediately.'

    b. *Watasi   wa   [soko   ni   hon   ga   at te]   hosi i.*
        1.sg   TOP there LOC book NOM be GER want-PRS
        'I want the book to be there.'

When *hosii* takes a control clause as its complement clause, however, it is not possible to emedd a non self controllable verb, as in (ii).

(ii) **Watasi   wa   gohan   ni   [PRO   suguni   deki te]   hosi i.*
     1.sg   TOP rice   DAT   immediately make GER want-PRS
     'I want the meal to be ready immediately.'

In (ii), the dative argument serves as a controller that controls PRO. The unacceptability of (ii) comes from the 'self controllability' condition imposed on control constructions.

(24) a. *Kuruma ga toori-sugi-ru.* b. *Ken ga gohan o tabe-sugi-ru.*
       car   NOM   go-pass-PRS   Ken   NOM   rice   ACC   eat-exceed-PRS
       'Cars pass by.'   'Ken eats too much rice.'

The complex predicate *toori-sugiru* 'pass by' in (24a), where *sugiru* carries the sense of 'pass', is a lexical compound which does not have a constituent structure transparent to the syntax, but *tabe-sugiru* 'eat too much' is a syntactic compound with a transparent syntactic structure. The *soo suru* 'do so' replacement test in (25), which is often used to assess the syntactic transparency of verbal constituents (Kageyama 1993), indicates that the two kinds of compound verbs indeed have distinct constituent structures.

(25) a. *Ken ga gohan o tabe-sugi-ta.*   *Mari mo soo si-sugi-ta.*
       Ken   NOM   rice   ACC   eat-exceed-PST   Mari   also   so   do-exceed-PST
       'Ken ate rice too much. Mari did so, too (=ate rice too much).'

     b. *Ken ga toori-sugi-ta. #Mari mo soo si-sugi-ta.*
       Ken   NOM   go-pass-PST   Mari   also   so   do-pass-PST
       'Ken passed by. Mari did so, too (≠passed by).'

The difference is further corroborated by the fact that *sugiru* expressing the sense of 'exceed' can be productively combined with any type of verb, while the verb *sugiru*, which carries the meaning of 'pass', cannot.

     When the excessive *sugiru* 'exceed' is combined with a verb, the entire complex is often regarded as forming a syntactic V-V compound (Kageyama 1993; Yumoto 2005; Kishimoto 2009). Nevertheless, *sugiru* possesses properties different from those of other verbs appearing in syntactic compounds (e.g. *kakeru* 'start', *naosu* 'repeat'). As shown in (26), *sugiru* 'exceed' can combine with a verb, a noun, an adjective, an adjectival noun, or a negated verb to form a complex compound, but the verb *naosu* 'repeat' can be directly combined with a verb only.

(26) a. *Kare wa {tabe/kodomo/sizuka/isogasi/sira-na}-sugi-ru.*
       he   TOP   {eat/child/quiet/busy/know-NEG}-exceed-PRS
       'He {eats too much/is too childish/is too quiet/is too busy/knows too little}.'

     b. *Kare wa {tabe/*kodomo/*sizuka/*isogasi}-naosi-ta.*
       he   TOP   {eat/child/quiet/busy}-repeat-PST
       'He repeated {the act of eating/*child/*quiet/*busy}.'

Arguably, *sugiru* is the only verbal predicate (or to be more precise, the only verbally-inflecting predicate) that can be combined with elements other than verbs to give rise to syntactic compound predicates. Importantly, the excessive *sugiru* forming a syntactically analyzable compound with a verb, a noun, or an adjective is devoid of its verbal property, i.e. it does not act like a verb. This is evidenced by the fact that it cannot be nominalized by *kata*-suffixation.

(27) a. *ryoori   no   tabe-{kake/naosi}-kata*
      dish   GEN   walk-{start/repair}-way
      'the way of {starting to eat the dishes/eating the dishes again}'

   b. *gohan   no   tabe-sase-kata*
      rice   GEN   eat-CAUS-way
      'the way of making (someone) eat rice'

   c. *kare   no   nagu-rare-kata*
      he   GEN   hit-PASS-way
      'the way of his being hit'

   d. **kare   no   {kodomo/karu/tabe}-sugi-kata*
      he   GEN   {child/light/eat}-exceed-way
      'the way of his {being too childish/being too light/eating too much}'

Since the nominalizing suffix *-kata* 'way' can be attached to a verbal element, compound verbs can be nominalized with the addition of the suffix *-kata*, as in (27a). The causative and passive suffixes *sase* and *rare*, which can be labeled as [+V, +Pred, +Aggl], can also host this suffix, as in (27b) and (27c), showing that the morphological status of bound versus free form does not affect the possibility of *kata*-suffixation. Nevertheless, V-*sugiru* compounds cannot be nominalized with the suffix *-kata*, as in (27d). In this connection, observe that *kata*-nominalization is permitted if *sugiru* is not part of a syntactic compound, as in (28).

(28) a. *(kuruma no)   toori-sugi-kata*   b. *zikan   no   sugi-kata*
      car   GEN   go-pass-way   time   GEN   pass-way
      'the way of cars' passing by'   'the way of time passing'

   c. *hatugen   no   do   no   sugi-kata*
      statement GEN   degree GEN   exceed-way
      'the way of making too extreme statements'

The fact that V-*sugiru* compounds in (27d) resist *kata*-nominalization suggests that the morphological dependent *sugiru* appearing in syntactic compounds does not function as a lexical verb, having the status as [−Pred, +Aggl]. In the light of the data (27), it is reasonable to state that the predicate *sugiru* forming a syntactic compound verb has undergone grammaticalization, and now serve as a functional element.

    In this section, we have discussed the issue of a correlation between the classification of lexical categories and their morphological forms. In Japanese, auxiliaries following a main predicate are formed into a morphologically tight unit via agglutination. On the basis of certain syntactic criteria, it has been seen that auxiliaries fall into either lexical or functional categories, which suggests that the categorial status of predicative elements does not strictly correlate with a morphological distinction

between "agglutinated" versus "non-agglutinated" forms. This in turn shows that morphological criteria are not necessarily reliable for the purpose of drawing a line between lexical and functional categories in Japanese.

4 Lexical categories from the cognitive-typological perspective

Cognitive linguistics places a focus on the nature and functional motivations of linguistic categorization. Linguistic typology (or simply, typology) examines attested patterns of variation across languages. This section will present an analysis of lexical categories in Japanese representing the collaboration of the two approaches.

4.1 Basics in the cognitive-typological approach to categorization

Usage-based cognitive linguistic investigations have repeatedly found that gradience is observed, when examined without contrary predilection, in almost every aspect of linguistic phenomena. In analyzing lexical category structures and organizations, cognitive approaches thus do not assume, a priori, "classical models" of categories, in which categories are discrete and phenomena are either inside of or outside of a given category (Lakoff 1987; Langacker 1987). Gradience, or lack of discreteness, however, does not mean chaos without any generalizations. It often implies "prototype" motivations in category organization (Rosch 1978; Taylor 1989). A prototype is a privileged subset of members of a category that represent the best exemplars of the category. The theoretical concept of a prototype is often relevant to and shared by typological approaches to linguistic categories.

Typological or cross-linguistic approaches are characterized by their acceptance of the fact of linguistic diversity that linguistic categories of particular languages can be irreducibly language-particular. Croft (1991, 2001) presents a theory based on a prototype approach to cross-linguistic grammatical patterns and has developed a universal definition (in the form of markedness patterns in typology) of the major lexical categories that is not constrained by the peculiarities of a particular language. To see how this theory can be applied in typologically characterizing lexical categories in Japanese, Uehara (1998) is introduced here. To indicate those points of departure of the cognitive-typological approach from the previous traditional and generative approaches, Uehara (1998) surveyed formal and structural criteria used to identify the so-called five "major" lexical categories of Japanese in eight past analyses, all of which explicitly discuss their grammatical behaviors (Hashimoto 1948; Kuno 1973; Martin 1975; Teramura 1982; Kageyama 1982; Miyagawa 1987; Shibatani 1990; Ohkado 1991). Having pointed out some problems arising

from language-particular aspects of their grammatical behaviors, Uehara discusses what contributions the cognitive-typological theory can make. Specifically, he adopts Croft's treatment of pragmatic (propositional) functions and reinterprets the formal and structural criteria used in the previous methods in cognitive-typological terms to provide a typological characterization of the lexical category structure of Japanese.

The five categories investigated as the "major" lexical categories in Japanese are Noun, Adjectival Noun, Adjective, Verbal Noun, and Verb,[5] and capitalization is used throughout this section to indicate that these categories in question are language-specific categories assessed in terms of some language-internal properties. These five categories are assumed to all be lexical categories with a relatively large membership, more or less distinct from one another in terms of their linguistic structural and/or semantic properties. They are termed the five "large" lexical categories (compared with words in minor lexical categories like *doodoo* 'dignified' or *onazi* 'identical, the same') in this section, to save the title "major" until they, or some other lexical categories, are assessed to be appropriate for it in cross-linguistic terms.

As one of the problems arising from language-specific aspects of lexical categories, Uehara (1998) points out that the criteria for defining these basic lexical categories are based on concepts which themselves are not well defined. Defining some language-specific properties of the Japanese language in terms of the concepts applicable to, say, English, does not secure sound characterizations of them, but rather brings about confusing situations for descriptive purposes. This is observed in apparent contradictions among the criteria proposed for the same AN category membership. One such pair is shown in (29). Each criterion is followed by those works that propose it.

(29) a. ANs can inflect (-*na* prenominally, -*da* sentence-finally).
          [Hashimoto 1948, Kageyama 1982]

      b. ANs do not inflect and need a copula (e.g. *da*), and take *na* before N.
          [Kuno 1973, Teramura 1982, Martin 1975, Miyagawa 1987, Shibatani 1990]

With the same criterion of inflection, ANs are inflectional according to some (notice the use of hyphens) in (29a) while ANs are non-inflectional in others (notice the same forms without hyphens) in (29b). This suggests that it is necessary to know what kind of grammatical behavior in Japanese each analysis treats as "inflection" before inflection can be used as a criterion for category membership of the language. In fact, it is of prime importance in characterizing the overall structure of lexical

---

5 These category labels are used throughout this chapter to be consistent, but different labels (not to mention their names in Japanese) are found for the same categories in different analyses. For instance, the term "Nominal Adjective" is used in Kuno (1973) and its Japanese close equivalent "meiyōshi" in Teramura (1982) for "Adjectival Noun" here.

categories in Japanese to examine what counts as inflection in Japanese and define it in terms of some intra-linguistic facts, which is done in the next section.

The other, more fundamental problem found in the previous approaches to lexical categories in Japanese, according to Uehara (1998), is the lack of a principled basis of choosing the criteria for lexical categorization. The criteria for a single category of Adjectival Nouns, for example, are various and numerous; even when the same criteria are collapsed together there still remain 19 different criteria by eight analyses. This lack of a principled basis virtually allows a scholar to privilege certain preferred criteria, whose data are hard to interpret typologically. This part of the section will briefly illustrate this and explicate how Croft's (1991, 2001) theory helps one to reinterpret them in cross-linguistic terms.

After presenting his data as to the inflection criterion (see (29b) above) for his argument for a certain lexical feature (cf. Section 3) for ANs in Japanese, Miyagawa (1987) presents, as another piece of evidence, the data reproduced in (30), to show that "the conditional *(ke)reba* attaches to V and A, but not to AN and N" (pp. 43–45):

(30) a. AN **sizuka-reba*   'if quiet'

     b. N   **sensei-reba*   'if a teacher'

     c. A   *utukusi-ke-reba*   'if beautiful'

     d. V   *tabe-reba*   'if (you) eat'

This part of his data can illustrate two accounts on which typological considerations can be implemented. The first concerns the theory of typological markedness (Greenberg 1966). By simply mentioning the form of the conditional as "*(ke)reba*", Miyagawa dismisses a structural difference between As and Vs, which can be observed in his data in (30), that *ke-reba* is for As while *reba* is used for Vs. In other words, for the forms indicating the same function of conditional, As always require an extra marking *ke* to the form for Vs. In the typological markedness terms, Vs are structurally "unmarked" (or "the least marked", in case there are more than two types) for conditionals while As are structurally "marked" (or "more marked" in comparison). This markedness criterion is crucial in linguistic typology for discovering cross-linguistic patterns, and is used in cross-linguistic approaches to lexical categories as well, as will be seen for those in Japanese later (see Croft 2001 for discussion of the other, "behavioral potential" markedness criterion, which is also used in typology but is omitted here due to space constraints).

The other point concerns the principle governing the selection of criteria for lexical categorization in cross-linguistic studies. Although Miyagawa's criterion above says "the conditional *(ke)reba* does not attach to AN and N" as seen in (30), it does not mean that they never take conditionals. They do, and the conditional form of the copula *nara(ba)* is used with them instead. So Miyagawa's paradigm in

(30) above could be supplemented with the data in (31) below for a descriptively fuller picture for cross-linguistic comparison.

(31)   a.   AN   *sizuka nara(ba)*   [quiet COP]   'if quiet' (cf. 30a)

       b.   N   *sensei nara(ba)*   [teacher COP]   'if a teacher' (cf. 30b)

What this criterion of conditionals by Miyagawa boils down to, then, is the same as his first one in (29b): whether some lexical item inflects itself or needs a copula in predication. Interestingly, this criterion as to the forms of lexical items in predication is among those criteria most commonly used (often in disguise) by the previous analyses surveyed, and at the same time it is one of the criteria used in Croft's cross-linguistic approaches to major lexical categories as well. The difference between the two, most naturally, is that some specific forms of the language (e.g. *(ke)reba* in (30) above) are used in the former, while in the latter such criteria are phrased in general, cross-linguistically applicable terms. Thus the criterion in question is rendered in the cognitive-typological approach as the (structural) markedness criterion of the lexical roots in the constructions indicating the pragmatic function of "predication".

    In addition to predication, the cognitive-typological approach employs "reference" and "modification" as the pragmatic functions directly relevant to the cross-linguistic identification of major lexical categories. Again, constructions indicating reference and modification are found among the groups of structural criteria commonly used and/or (often implicitly) repeated in the surveyed previous analyses of the five large lexical categories in Japanese. They are of course expressed in more or less language-specific ways, such as "can function as subject (or object)" or "can take case particles" for reference and "take *na* prenominally" (as in (29) above) for modification.

    The pragmatic function as the basis of major category definitions is a propositional act, and it is thus referred to as "propositional" function hereafter in this chapter. A propositional act is analogous to an illocutionary speech act, but it is a speech act that structures information inside the proposition rather than modifying the proposition as a whole (Searle 1969). In performing a speech act, the speaker must perform a series of propositional acts. The most important of these are reference, predication (since these are required for every proposition), and (to a lesser extent) modification. The reference function is defined as "to get the hearer to identify an entity as what the speaker is talking *about*" (Croft 1991: 52). The predication function is to say "what the speaker intends to say about what he is talking about (the referent)" (Croft 1991: 52) to ascribe something to it, and the modification function is "to enrich the nominal image by an additional feature" (Croft 1991: 123).

    Croft's (1991) cross-linguistic study on lexical categories found some patterns repeated across languages in the interaction of propositional function and semantic class. The semantic classes of objects, actions, and properties (taken originally from

the traditional approach) are only a small subset of the semantic classes of words found in human languages, and are defined in terms of four semantic properties of valency, stativity, persistence, and gradability (see Croft 1991). It is found that there is a prototypical correlation between the semantic class of object and the function of reference so that words denoting objects cross-linguistically take the prototypical markedness pattern for nouns, in which their lexical roots are the least marked for reference. A similar prototypical markedness pattern holds between actions and predication for verbs and between properties and modification for adjectives.

Such a prototypical correlation is most conspicuously revealed and illustrated in the structural markedness patterns in English shown in Table 3 below, slightly modified from Croft (1991:67) with English examples from Croft (1991:53):

Table 3: Typological correlations with English examples (from Croft 1991)

	Reference	Modification	Predication
Objects	UNMARKED NOUNS (vehicle)	genitive, PP's on nouns (vehicle's, *of/in*/etc. the vehicle)	predicate nominals (*be* a/the vehicle)
Properties	deadjectival nouns (tall*ness*)	UNMARKED ADJECTIVES (tall)	predicate adjectives (*be* tall)
Actions	action nominals, complements, infinitives, gerunds (destruc*tion*, *to* destroy)	participles, relative clauses (destroy*ing*, destroy*ed*)	UNMARKED VERBS (destroy)

4.2 Toward a formal definition of *katsuyō*-inflection in Japanese

This section addresses a basic, language-specific question postponed earlier: what does it mean to say that a word *inflects* in Japanese? The word *inflection* is a translation of "katsuyō" in Japanese, a traditional grammatical term. Past analyses by and large agree that Nouns do not inflect while Adjectives and Verbs inflect in Japanese. As for words in the AN category, however, as we saw above in the survey of past analyses in (29), there are two competing positions: one posits that ANs inflect and the other that ANs do not. This is illustrated below by the most basic (i.e. the present, affirmative, declarative) forms[6] of representative examples of each

---

6 This form is attested by hearsay evidential markers, *soo da* 'it is said that' and *to iu koto da* 'it is the case that'.

(i)   *Kirei( )* <u>*da*</u>   *{soo da/to iu koto da}.*
      pretty COP {it.is.said.that/it.is.the.case.that}
      '{It is said that/It is the case that} (it) is pretty.'

Also, the unmarked form prototypically appears before prototypical conjunctive clause particles like *kedo* 'although', and *kara* 'because'.

(ii)   *kirei( )* <u>*da*</u>   *{kedo/kara}* …
      pretty COP {although/because}
      '{Although/Because} (it) is pretty, …'

category, where the use of hyphen indicates the inflectional status (note the zero pronominal nature of the Japanese language, where complement nouns can be implicit and predicates alone can form sentences, indicated by the capitalization of the first letter, in the example sentences in (32) and below):

(32) a. N   *Hon da.*   [book COP]   '(It) is a book.'

b. AN *Kirei(-)da.*   [pretty-INFL/COP]   '(It) is pretty.'

c. A   *Huru-i.*   [old-INFL]   '(It) is old.'

d. V   *Tabe-ru.*   [eat-INFL]   '(It) eats (it).'

This disagreement as to the inflectional status of ANs, together with the controversy as to whether to establish a separate category of ANs, is observed not only in the survey above, but throughout the history of syntactic category analyses of Japanese, which dates back to FUJITANI Nariakira (1778). Fujitani first recognized ANs as a category called *ari-sama*, which is a subcategory of *sama*, which corresponds roughly to the Adjective category. Since then, two prominent traditional grammarians, Hashimoto and Tokieda, for example, took opposite sides (see Mizutani 1951 for a summary), and the controversy has continued up to the generative approaches (Kageyama 1982; Miyagawa 1987) as well, as seen above in (29). Therefore, what counts as *katsuyō*-inflection needs to be made clear before it can be used as a criterion in any lexical category analyses of Japanese.

The important question to ask is whether there is any linguistically salient contrast among the four kinds of predicates in (32). Uehara (1998) answers in the affirmative to this question, and demonstrates that the contrast between N and AN on the one hand and A and V on the other is a very salient one. The contrast shows up when one starts looking at the forms of predications with more marked functions than the basic, unmarked ones. Among notable instances of contrast is the reduplication for emphasis of predicates, which are shown in Table 4.

Table 4: Lexical roots in unmarked and emphatic predicates

	Unmarked predicates		Reduplication for emphasis	
N	*Hon da.*	'(It) is a book.'	*Hora! Hon, hon!*	'See! (It) IS a book!'
AN	*Kirei da.*	'(It) is pretty.'	*Hora! Kirei, kirei!*	'See! (It) IS pretty!'
A	*Huru i.*	'(It) is old.'	*Hora! *Huru, huru! => Huru i, huru i!*	'See! (It) IS old!'
V	*Tabe ru.*	'(It) eats (it).'	*Hora! *Tabe, tabe! => Tabe ru, tabe ru!*	'See! (It) DOES eat (it)!'

One contrast observed above is that the first elements (i.e., lexical roots) of N and AN predicates can stand alone in the emphatic predicative function while those of A and V predicates cannot. That is to say, A and V roots are "bound" morphemes unlike N and AN roots.

The same can be observed also in question predicates, direct or indirect, most of which involve the question particle *ka*. Table 5 below shows the most obvious case of contrast, in which in casual colloquial speech the lexical roots are used in questions without the question particle *ka* but with the question intonation.

Table 5: Lexical roots in unmarked and interrogative functions

	Unmarked predicates		Predicates in casual-style question	
N	*Hon* da.	'(It) is a book.'	*Hon* ?	'Is (it) a book?'
AN	*Kirei* da.	'(It) is pretty.'	*Kirei* ?	'Is (it) pretty?'
A	*Huru* i.	'(It) is old.'	**Huru* ? => *Huru* i ?	'Is (it) old?'
V	*Tabe* ru.	'(It) eats (it).'	**Tabe* ? => *Tabe* ru ?	'Does (it) eat (it)?'

The same pattern of contrast can be observed in the use of lexical roots with the question particle *ka*, with *ka sira* 'I wonder if . . . .' (mostly by female speakers) and with *ka doo ka* 'whether or not. . . .' (indirect/embedded questions), only the first of which is shown below in (33) due to space limitations:

(33) a. N *Hon ka?*   'Is (it) a book?'

     b. AN *Kirei ka?*   'Is (it) pretty?'

     c. A **Huru ka? => Huru i ka?*   'Is (it) old?'

     d. V **Tabe ka? => Tabe ru ka?* 'Does (it) eat (it)?'

The above examples demonstrate that the second elements of A and V predicates, *i* and *ru* (*u* in case of consonant-root Vs) respectively, are more tightly joined with their first element (i.e. "root"), than are those of N and AN predicates (i.e. *da*).

The other group of predicates that exhibit the same pattern of contrast are epistemic modality predicates, which are, again, predications with marked functions. They include *daroo* 'it is probable that', *desyoo* 'it is probable that (polite)', *ka mo sirenai* 'may be', *ni tigainai* 'surely, must be', and *mitai* (*da*) 'seem like'. Only the pattern with the lexical roots with *daroo* 'probable' is shown in (34) due to space limitations, but the same pattern holds with the other modality markers as well:

(34) a. N *Hon daroo.*   '(It) is a book, probably.'

     b. AN *Kirei daroo.*   '(It) is pretty, probably.'

     c. A **Huru daroo. => Huru i daroo.*   '(It) is old, probably.'

     d. V **Tabe daroo. => Tabe ru daroo.* '(It) eats (it), probably.'

All the data presented here demonstrate that though all the basic predicate types in (32) seemingly are composed of two elements, there is a salient difference between N

and AN predicates on the one hand and A and V predicates on the other in the con-nectedness between the two elements of their predicates. In other words, the roots of A and V unmarked predicates are more "bound" to their following elements than those of N and AN predicates. With this contrast in mind, one can now safely use the "-" notation as in (35) below:

(35)   a.   N   *Hon da.*   [book COP]   '(It) is a book.'
       b.   AN   *Kirei da.*   [pretty COP]   '(It) is pretty.'
       c.   A   *Huru-i.*   [old-INFL]   '(It) is old.'
       d.   V   *Tabe-ru.*   [eat-INFL]   '(It) eats (it).'

With this distinction in boundness as the language-internal basis for deter-mining inflectional categories (A and V) and non-inflectional ones (N and AN) in Japanese, one may now proceed to typological characterization of the lexical cate-gories in Japanese in the next section. However, there are two points worth noting here about this boundness distinction of the language, which will have direct relevance to our later discussions. These two points are discussed in the next two subsections, respectively.

4.2.1 Gradience in boundness

In the previous section we have focused on the salient contrast in boundness of lexical roots between N and AN on one hand and A and V on the other (see (35) above). However, this by no means implies that the morphological bound/free dis-tinction itself is discrete. Instead, as will be seen below, the degree of boundness varies even among the members of inflectional categories (of bound lexical roots) and of the other, non-inflectional categories.

Both lexical roots of A (e.g. *huru-i* 'old') and V (e.g. *tabe-ru* 'go') are bound (i.e. 'not free') compared with those of Ns and ANs, in that their inflectional ending (*-i* and *-(r)u*, respectively) cannot be dropped in the context where the lexical roots of the latter can stand alone. A difference in the degree of boundness, however, shows up between the two inflectional lexical categories and indicates that A roots are the freer of the two. This is shown by the fact that the so-called "Adjectival conjugational ending drop construction" (Konno 2012) is available, which expresses "the speaker's immediate reaction to a given situation in which he/she is involved at the time of utterance" (Konno 2012). In that construction, as its name indicates, A roots are used without their inflectional ending *-i*. No construction with the same function is available for V roots, and even when occurring in similar functions they always require their ending *−(r)u*. This contrast is shown in (36) below:

(36) a. A *Huru!* 'Old!' (<= *huru-i*)

    b. A *Dasa!* 'Uncool!' (<= *dasa-i*)

    c. A *Kimoti-waru!* 'Disgusting!' (<= *kimoti-waru-i*)

    d. V *Wakar*(-u)!* '(It) makes sense!' (*wakar-u* 'understand')

    e. V *Simi*(-ru)!* '(My teeth) aches!' (*simi-ru* 'soak through')

It should be also noted that a phonological restriction renders many V roots inherently bound, unlike A roots. The lexical roots of Verbs may end with consonants (as in *k* of *ik-u* 'go' and *r* of *wakar-u* 'understand') or with vowels *i* or *e* (as in *oki-ru* 'arise' and *tabe-ru* 'eat'). The mora structure of Japanese does not allow syllable-final, non-moraic consonants. In other words, no lexical roots of consonant root Verbs can stand alone without their inflectional ending.

    The existence of the inflectional ending drop construction for A but not for V, together with the language's phonological restriction on the V (consonant) roots, suggest that the degrees of boundness of the bound A and V lexical roots can be shown as A > V, where the left side of > is freer.

    Moving on to the other side of the boundness distinction, namely, N and AN, one sees some similar difference in the degree of boundness between the two non-inflectional categories. This difference in boundness can be observed in their occurrences with case-marking particles. N roots can be directly followed by case-marking particles, while AN roots cannot and require a nominalizing suffix *sa* just like A roots, as in (37), where the notation **(sa)* means that *sa* is obligatory in this context.

(37) a. N *hon {ga/o}* [book {NOM/ACC}]

    b. AN *kirei-*(sa) {ga/o}* [pretty-*(ness) {NOM/ACC}]

    c. A *huru-*(sa) {ga/o}* [old-*(ness) {NOM/ACC}]

This indicates that between the two non-inflectional categories of free lexical roots, N roots are the freer of the two (N > AN).

    Table 6 below summarizes the above discussion:

Table 6: Degrees in boundness of lexical roots

	inflectional ending drop construction	marked forms of predication	directly with case-marking particles
N	(ok)	ok	ok
AN	(ok)	ok	*
A	ok	*	*
V	*	*	*

Table 6 clearly demonstrates that the bound/free distinction is a matter of degree and the degrees in boundness of lexical roots of the four lexical categories can

be shown in the form of a cline in (38) below (functional motivations for the cline will be discussed later):

(38)   N > AN > A > V

This boundness cline provides a possible explanation of why some researchers, traditional and generative alike, postulate AN to be inflectional along with A and V: they assign the boundness level between N and AN for the non-inflectional/inflectional division, while others assign the one between AN and A.

As discussed in Section 4.2, however, the current analysis concurs with the latter (i.e. AN, as well as N, is non-inflectional). This is because the *katsuyō*-inflection in question is concerned with predication function, while the use with case-marking particles indicating the boundness level between N and AN is concerned with reference rather than predication (see the discussion above).

4.2.2 The boundness distinction as a cardinal formal property of Japanese

This contrast in boundness of lexical categories is such a characteristic property of the predicate structure of the language that it captures a generalization, and thus facilitates the description, of not only all the predicate patterns enumerated in Section 4.2, but also other structural patterns of the language, of which only two are noted here.

The first represents the interaction of lexical categories with the pragmatics of the language found in the difference between default/masculine and feminine styles in colloquial speech. The following in Table 7 is an illustration of the difference, by using a sequence of a predicate + a sentence final particle of assertion, *yo* 'I tell you':

Table 7: The boundness distinction and a style difference

	Default/Masculine style	Feminine style	meaning
N	*Hon <u>da</u> yo.*	*Hon yo.*	'(It) is a book, I tell you.'
AN	*Kirei <u>da</u> yo.*	*Kirei yo.*	'(It) is pretty, I tell you.'
A	*Huru i yo.*	*Huru i <u>wa</u> yo.*	'(It) is old, I tell you.'
V	*Tabe ru yo.*	*Tabe ru <u>wa</u> yo.*	'(It) eats (it), I tell you.'

Here also, we find one pattern for Nouns and Adjectival Nouns, and another for Adjectives and Verbs – a contrast which now we are able to capture by the epithets, "non-inflectional" and "inflectional" groups.

The other point concerns the productivity of lexical categories of the language: this boundness of the A and V roots is such that it has made these categories closed

classes while making the other, "free" (non-inflectional) roots the only source for new word coinage, even though cross-linguistically inflection does not always block new members and such major categories in other languages are often open classes. That is, when new terms are coined in Japanese, they become members of the A or V classes with great difficulty, and instead become members of open classes such as AN and N. A striking example of this pattern can be seen in the formation of loan words in the language. There are a large number of English adjectives that have been borrowed into the Japanese language, but they seldom, if ever, become Adjectives, but ANs. *Ritti* 'rich' (39a) and *ereganto* 'elegant' (39b) below are among such examples. In (39), the corresponding adnominal forms, which follow the AN but not the A pattern, are shown in square brackets.

(39) a. *Ano   hito   wa   ritti {da/*-i}.*   [*ritti {na/*-i}   hito*]
          that person   TOP   rich   rich   person
          'That person is rich.'   'a rich person'

     b. *Hun'iki   ga   ereganto {da/*-i}.*   [*ereganto {na/*-i}   hun'iki*]
          atmosphere   NOM   elegant   elegant   atmosphere
          'The atmosphere is elegant.'   'an elegant atmosphere'

This is true also of English verbs borrowed into Japanese. They follow the very productive Verbal Noun pattern, i.e. VN + *suru*. The following are only some of the numerous instances of VN words from English[7]:

(40) *kopii suru,   katto suru,   anaunsu   suru,   hitto suru*
       copy   do   cut   do   announce do   hit   do
       'to copy'   'to cut'   'to announce'   'to make a hit'

In fact, the boundness of A and V roots has been a characteristic of the Japanese language throughout its history, so that one does not have to wait for the recent borrowings from English for the evidence of their closed class status. In its long history of contact with the Japanese language, Chinese vocabulary has flowed into the language, and loan words from Chinese constitute a major part of the vocabulary of the Japanese language, most of which are Ns, ANs and VNs (see Uehara (2003) for a scenario of the historical development of ANs into major category in present-day Japanese).

---

7 There is another less productive, but still useful colloquial pattern of N+*r u* (e.g. *kopir u* 'to copy' from *kopii* 'copy' and *memor u* 'to take notes' from *memo* 'notes'). See Uehara (1998: Chap. 4) for more examples.

(41) a. AN *kantan na,* *kyodai na,* *kimyoo na,* *yuudai na*
          'simple'   'gigantic'   'strange'   'magnificent'

     b. VN *soodan suru,* *syoyuu suru,* *baisyuu suru,* *kenkyuu suru*
          'to consult'   'to possess'   'to purchase' 'to research'

Thus, the existence of the two categories that are often noted as categories "unique" to the Japanese language (Shibatani 1990), can be ascribed to this boundness of the two inflectional categories, Adjectives and Verbs.

4.3 Typological characterization of lexical categories in Japanese

Now that what counts as inflection in Japanese is clearly laid out, the structural criteria in Japanese for the major propositional functions of reference, modification and predication can be summarized as follows:

Reference: whether or not it takes nonzero morpheme(s) such as *sa* '-ness' when used with case-marking particles

Modification: whether or not it takes nonzero morpheme(s) such as *no* or *na* before N

Predication: whether it inflects, or does not inflect and requires non-zero morpheme(s) such as the copula or the dummy verb in the predicate

4.3.1 The five "large" lexical categories

The criteria listed above are used to typologically characterize the structural properties of the five lexical categories in Japanese, which are summarized in Table 8 below (the least marked forms (present, affirmative indicative) are listed):

Table 8: Structural properties of five "large" lexical categories in Japanese

	Reference (take case particles)	Modification (before N)	Predication (inflect)
N	ROOT	ROOT + *no*	ROOT + *da*
VN	ROOT (+*suru* + *koto/no*)	ROOT + *no/suru*	ROOT + *da/suru*
AN	ROOT *sa* '-ness'	ROOT + *na*	ROOT + *da*
A	ROOT *sa* '-ness'	ROOT *i*	ROOT *i*
V	ROOT *(r)u* + *koto/no*[8]	ROOT *(r)u*	ROOT *(r)u*

---

8 Some may argue that the stem form (root *(i)*) of the Verbs (e.g. *tutum i* 'package' from *tutum u* 'to pack') be in the V reference cell in Table 8. Refer to Uehara (1998: Chap. 2), which points out that the stem forms are often bound (e.g. **tabe* from *tabe ru* 'to eat' and **k i* from *k uru* 'to come') and that when they are free, they typically refer to some objects involved in the actions (e.g. *tutum i* above) or some aspects of them (e.g. *kaer i* 'the time/occasion of one's returning' from *kaer u* 'to return'), rather than to the actions themselves, which are referred to by the form in the cell in question in Table 8 (i.e., ROOT *(r)u* + *koto/no*).

The above table of the structural properties of the five "large" lexical categories in Japanese demonstrates two respects in which propositional functions interact with Japanese lexical category organization. First, Table 8 shows no difference in structural markedness between modification and predication,[9] which indicates that modification has not played a major role in the category organization of the language in addition to what predication has already. This can be considered to be a reflection of the fact that the modification function is optional and secondary compared with the two major propositional functions, both of which are required in every proposition. Present-day Japanese thus can be characterized as one of the languages in which this secondary status of modification is structurally observed. (This point and its relevance to the non-prototypical statuses of A and AN in the language will be discussed later.) Second, following from the first, reference and predication have played major roles in the language's lexical categorization: Noun, at the top of the list of large categories in Table 8, is the least marked for reference and others are progressively more marked as they go down the list. The markedness order is reversed for predication, for which Verb at the bottom is the least marked. In other words, the boundness degree order in (38), which is the determining factor of the order of paradigms in Table 8, is motivated by the propositional functions of predication and reference. The most "bound" V roots and their structural paradigm are motivated by predication function, while the freest N roots and their paradigm are by reference function. The other lexical roots and their paradigms of intermediate levels of boundness are in between in that regard.

From the comparison of Table 8 with Table 3, where Croft's cross-linguistically attested, prototypical lexical category paradigms are exemplified with English noun, verb, and adjective, the following points regarding the five Japanese lexical categories are observable as their functional-typological characteristics (for detailed discussions of each point with concrete examples, refer to Uehara 1998):

(a) N exactly follows the prototypical nominal paradigm in the Objects row like English nouns, where the root is unmarked for reference but is more marked for modification and predication. VN is identical with N in its forms and markedness pattern, the only difference being the existence of an additional pattern with *suru* 'do' for VN. In other words, VNs are Ns with the additional ability to compound with *suru* into compound level Vs.

(b) Neither AN nor A in Japanese exactly follows the prototypical adjectival paradigm in the Properties row, which English adjectives follow: ANs are like English adjectives in reference and predication functions, but in modification, ANs differ from them and need an extra morpheme (like Ns). In that sense, ANs are structurally "nouny" adjectives (Wetzer 1992). On the other hand, Japanese As are like

---

9 This structure for the present day Japanese can be ascribed to the loss of the formal distinction between adnominal and predicative forms that used to exist for many Verbs and Adjectives in Old Japanese. See the relevant discussion in Uehara (1998).

English adjectives only in modification (the least marked form of the inflection). Unlike English adjectives, Japanese As (like Vs) can stand by themselves (i.e., without taking the copula *da*) in predication. In that sense, As are "verby" adjectives (Wetzer 1992). As are as unmarked for reference as for modification.

(c) Except for modification, V in Japanese follows the prototypical verbal paradigm in the Actions row, which English verbs follow: Vs thus resemble English verbs in predication (unmarked) and in reference (more marked), but, unlike them, Japanese Vs are unmarked also in modification without requiring any extra morpheme (i.e., the same form as in predication).

Table 8 furthermore makes possible some observations regarding the way the five categories are inter-categorically related and, accordingly, how they are organized overall.

Verbal Noun constitutes a non-prototypical subcategory of Noun: VNs belong to the Noun category because they possess the same structural paradigm with Ns, and they are less-prototypical members of the Noun category because VNs in the additional structural pattern (ROOT + *suru*) are most marked for reference (ROOT + *suru* + *koto/no*). Adjectival Noun sides with N (rather than A): AN and N share the same markedness pattern (nonzero morphemes) in both modification and predication (and in the latter they even share the same form *da*). Thus, AN and N together can form a super-ordinate category in contrast with A and V, which are unmarked for both modification and predication. In the super-ordinate category subsuming N and AN, Ns are prototypical members while ANs are less prototypical ones, since only Ns are unmarked for reference.

Adjective and Verb present no contrast in structural markedness (both inflecting themselves without requiring any extra morpheme) with respect to predication and modification functions. In reference only, the two differ and V is the more marked. A and V in Japanese, unlike adjectives and verbs in English, are thus structurally very similar to each other, which suggests that they together form a superordinate category of which the two are subcategories, as opposed to the other three categories.

Thus, simply by looking at the structural characterizations of the five large lexical categories, we find the two super-ordinate categories of N (subsuming VN) and AN on the one hand, and V and A on the other. One may notice this major division corresponds to the non-inflectional/inflectional distinction based on the boundness as discussed in the previous sections. In other words, the boundness distinction constitutes the first "cut" into the lexical category structure of Japanese, thus characterizing it as the language with two, not three, "major" categories, of which the five large lexical categories are sub-classifications. It should be pointed out in passing, however, that this situation of Japanese is typologically not rare, where adjectives do not structurally constitute an independent category of its own, but become a subcategory of other major categories, and that it is not totally unmotivated considering the secondary status of modification compared to predication

and reference. In Chinese and Thai, for example, adjectives behave like 'stative verbs' (i.e. a subcategory of verbs), not requiring a copula in predication. In Korean, adjectives and verbs both inflect and share their inflectional endings (Uehara and Kumashiro 2007).

One of the two major lexical categories of Japanese subsuming the non-inflectional subcategories are termed the "Nominal" here, since N represents its prototype, following the cross-linguistic markedness pattern for nouns, while AN and VN represent extensions from N. The other major category subsuming the inflectional subcategories is termed the "Verbal" because Verb basically follows the prototypical pattern for verbs and represents its prototype.

The dichotomous Nominal/Verbal distinction draws the line between ANs and As in Table 8. Figure 1 below summarizes the discussions and schematically represents the overall lexical category structure of Japanese:
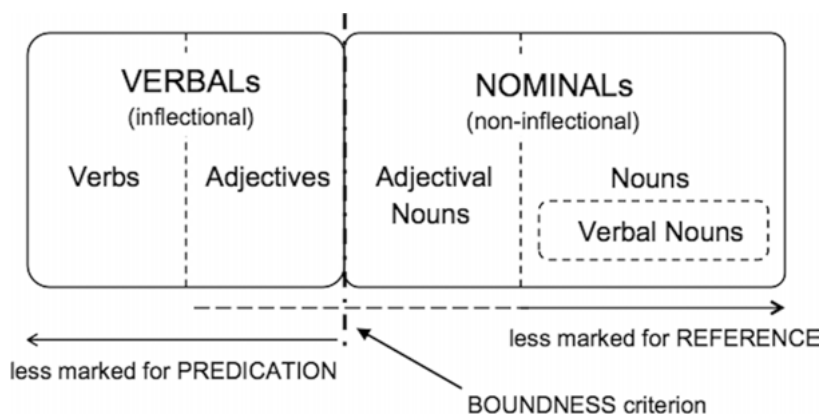

Figure 1: Lexical category organization in Japanese

4.3.2 Other approaches and other minor/non-prototypical subcategories

In the light of Table 8 and Figure 1, which capture typology-based structural characterizations of the 5 large lexical categories of Japanese and overall Japanese lexical category organization, respectively, one can readily assess various other approaches to them. They differ in which of the structural criteria are given greater weight than others, which semantic/functional criteria are additionally taken into consideration, and which granularity levels of distinctions are to be accounted for. Furthermore, with the 5 constructional paradigms in Table 8 as basic templates, one can also characterize other, minor lexical categories of the language as non-prototypical subcategories of them.

As noted in Section 2, Tokieda (1950) argues explicitly against setting up Adjectival Noun as an independent lexical category. But his argument, which leads to categorizing all ANs into his *taigen* category, the Nominal category in Figure 1, indicates that he weighs the commonalities between N and AN in the predication function over their differences in the reference and modification functions. In contrast, Hashimoto (1947) and Kageyama (1982), which describe AN as inflectional in (29a), attach no weight to the formal identity between AN and N in predication. Sakuma (1951) goes further to collapse A and AN together by giving more weight to semantic and functional, rather than structural, criteria, as indicated by the following quote from Sakuma (1951: 54): "it is reasonable to posit one lexical category by grouping Adjective and Adjectival Noun together by setting aside the morphological differences between them, because they resemble each other so completely in their meanings and functions (translation by SU)." He sets up a lexical category called *seijōgo*, which subsumes the two as its subcategories, and calls A the first *seijōgo* and AN the second *seijōgo*.

Some propose to take finer-grained distinctions into consideration to argue for an analysis that stands apart from the popular five major categories approach. One such analysis is found in Muraki (2012) and his previous works cited therein. Muraki's analysis can be characterized as an extension of Sakuma (1951) in that he posits the third "adjective" category on top of his first and second "adjective" categories, which respectively correspond to first and second *seijōgo* in Sakuma's approach and A and AN in Table 8 and Figure 1 above in the current approach. His third "adjective" takes *no* (instead of *na*) in modification and the copula *da* in predication, but is not used in reference (Muraki 2012: 149). He lists the following examples in their forms with *no*:[10]

(42) *sinku no* 'crimson-colored',   *batugun no* 'outstanding/unrivaled',

   gokaku no 'evenly-matched',   *marugosi no* 'unarmed',

   mayakasi no 'fake'

His third "adjective" can be characterized as a non-prototypical subcategory of the Noun category, lying on its border with the AN category, in the current approach: it takes the forms of the N paradigm in Table 8 except it lacks the form in reference. Muraki's classification of the group of words as the third "adjective" after A and AN as the first and second ones, respectively, reflects their structural properties: among his three "adjectives", they are the least different from N.

---

10 See also "precopular nouns" in Martin (1975). Words of Muraki's third "adjective" category listed here as well as many others such as *nama* 'raw' and *ippai* 'full' are discussed in Uehara (1998: Chap. 3) as those lexical items which are translated into adjectives in English but take *no* instead of *na* in modification.

Some other non-prototypical subcategories can be characterized as such, not by lacking a form or two of one paradigm, but by having forms of more than one paradigm, listed in Table 8. Uehara (1998: Ch. 3) discusses what he calls double-function cases of many words (typically denoting abstract concepts or human attributives) among ANs, such as *heiwa* 'peaceful/peace' and *kenkoo* 'healthy/health'. They are ANs (taking *na* in modification) but take *no* as well in modification and/or can take case-marking particles in reference as in (43), which indicates their ambivalent status between AN and N. Such ambivalent AN/N words constitute a non-prototypical subcategory of the Nominal category.

(43) a. **heiwa** *na*   *kuni*   b. *kono kuni*   *wa* **heiwa** *da.*
                 country   this country TOP   COP
          'a peaceful country' (AN)   'This country is peaceful.' (AN)

    c. **heiwa** *no* *sisya*   d. **heiwa** *ga*   *dai-iti*   *da*
                 messanger   NOM first   COP
          'a messenger of peace' (N)   'Peace comes first.' (N)

Also among non-prototypical members of the Nominal category are those words conventionalized with forms of the AN and VN paradigms, such as *manzoku na/suru* 'satisfactory/get satisfied' and *nangi na/suru* 'difficult/have difficulty' as in (44) below (See also the discussion of such Sino-words in Chapter 3 [Kobayashi, Yamashita and Kageyama, this volume]). In a similar vein, some lexical roots (e.g. *atataka* 'warm', *okasi* 'funny') can be characterized as having been conventionalized with forms of A and AN to various degrees, as in (45), where their forms in modification are followed by those in predication in square brackets (See also Uehara (1998:89) for the list of such words).

(44) a. **manzoku** *na*   *kotae*   b. *sono*   *kotae*   *ni*   **manzoku** *suru*
                  answer   the   answer   with
          'a satisfactory answer' (AN)   'be satisfied with the answer' (VN)

(45) a. **atataka** *{-i/na}* *hizasi*   [*hizasi*   *ga*   **atataka** *{-i/da}*]
                      sunlight   sunlight   NOM
          'warm sunlight' (A/AN)   'The sunlight is warm.' (A/AN)

    b. **okasi** *{-i/na}* *hanasi*   [*sono*   *hanasi* *wa*   **okasi** *{-i/*da}*]
                   story   the   story   TOP
          'a funny story' (A/AN)   'The story is funny.' (A)

The existence of such non-prototypical members/subcategories demonstrates that they have as their prototypes the five "large" lexical categories and their structural paradigms in Table 8. They are in turn prototypically organized as represented in Figure 1, where Noun and Verb are their prototypical cores, respectively motivated by propositional functions of reference and predication.

5 Conclusion and future research perspectives

In this chapter, some issues surrounding the Japanese lexical categories (parts of speech) have been discussed. In the literature on Japanese, words have been classified in a number of different ways, and the currently available proposals concerning their categorization differ significantly depending on what kind of perspective has been adopted. In traditional Japanese grammar, parts of speech are defined with an emphasis on their morphology. The point of divergence in traditional Japanese grammar lies in the morphological status of words, in relation to their categorization, and its major controversies concern the status of adjectival nouns and the treatment of minor categories (such as auxiliaries). On the other hand, the generative and cognitive-typological approaches aim to characterize categories from the perspective of language universals, although the former often takes them as a given while the latter motivates them by cross-linguistically attested data. The major difference between the two views lies in the fact that the generative approach takes lexical categories to be discrete, while the cognitive-typological approach does not a priori assume them to be, admitting of them being non-discrete entities, which often form a continuum.

In the generative framework, lexical categories are defined via cross-categorical features, which are claimed to be universal (i.e. they are regarded as valid across languages). In Japanese generative grammar, there have been issues on how the lexical categories should be defined in terms of cross-categorial features, since the language has more major lexical categories than can be defined by the simple feature system [±V, ±N], due to the presence of adjectival nouns and verbal nouns. Further, Japanese is an agglutinative language, where bound predicative elements (regardless of their categorial status) form a morphologically tight unit with the main predicate. Thus, some but not all predicative elements traditionally classified as dependent auxiliaries can be identified as major lexical categories. Besides, function words (belonging to functional categories) can be free or bound morphemes. These facts suggest that morphological criteria do not necessarily constitute reliable heuristics to distinguish between the major lexical and the functional categories of Japanese predicative elements.

In the cognitive-typological approach, words (and their categories) may be gradient, and the structural organization of five lexical categories in Japanese can be identified with reference to criteria utilized in a prototypical approach. Using cross-linguistically applicable markedness criteria, it has been shown that Japanese can be characterized as possessing two major lexical categories (inflectional and non-inflectional categories), which have the five lexical categories as their subcategories. The dividing line between the two major categories is the boundness criterion, which is a linguistically salient contrast in morphological boundness and is used to define the inflection of the language. Inflectional and non-inflectional

categories are unmarked for the two major pragmatic functions of predication and reference, respectively.

It is worthwhile to note that nouns and verbs are taken as constituting rudimentary categories in all the approaches, which is naturally expected, given that both are likely universal categories cross-linguistically (cf. Whaley 1997). By contrast, some minor categories, which do not appear to be common or universal typologically, give rise to a number of controversies. In this regard, Japanese morphology is not necessarily effective in evaluating the category memberships; traditionally, categories are distinguished placing emphasis on morphological criteria, but as noted earlier, there is no tight correlation between the morphological distinction of "agglutinated/bound" and "free" and the distinction of "lexical" versus "functional". In addition, lexical items sometimes seem to have gradient properties or could be ambiguous in their class membership, in which case category labels cannot be assigned uniquely.

There are many theoretical issues that are worth pursuing for future research. The most prominent and yet fundamental issue concerns the criteria for distinguishing categories. The criteria primarily used for this purpose differ among the three approaches. Traditional Japanese grammar gives precedence to morphological criteria over syntactic ones. Generative grammar uses both morphological and syntactic criteria equally to distinguish categories, but cognitive grammar adopts semantic and functional criteria as well as structural ones. Then, the question inevitably arises as to what would be the optimal way of determining the status of lexical categories. Furthermore, in the classification of lexical categories, super-categories (such as *taigen* (nominal) and *yōgen* (predicative), which groups verbs, adjectives and adjectival nouns together) and sub-categories (as seen in the finer classifications of particles) are often posited. This raises the question of what level of classification is appropriate for capturing basic generalizations in Japanese grammar, and it still remains to be seen how fine-grained criteria should be used to identify lexical categories. It would also be interesting to see whether the generalization on lexical categories in standard Japanese applies to various dialects of Japanese, since dialects sometimes show curious properties that help us clarify issues that are not readily resolved only by considering standard Japanese (see e.g. Kudō 2004).

Acknowledgments

collaborative research project "Syntactic, Semantic, and Morphological Characteristics of the Japanese Lexicon" (Project leader: Taro Kageyama, National Institute for Japanese Language and Linguistics, 2010–2015).

References

Abney, Steven. 1987. *The English noun phrase and its sentential aspect*. Cambridge, MA: MIT dissertation.

Backhouse, Anthony. 2009. On the status of *nai* in modern Japanese. *Journal of Japanese Linguistics* 25. 59–74.

Baker, Mark. 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Bloch, Bernard. 1946. Studies in colloquial Japanese I: Inflection. *Journal of the American Oriental Society* 66. 97–109.

Brinton, Laurel and Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.

Carnie, Andrew. 2007. *Syntax: A generative introduction*. 2nd edition. Malden, MA: Blackwell.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Emonds, Joseph. 1985. *A unified theory of syntactic categories*. Dordrecht: Foris.

Fujitani, Nariakira. 1778. *Ayui shō: Yosoinokata* [Suffixal dependents: Inflectionals]. In Masao Takeoka (ed.), 1961–1962. *Fujitani Nariakira zenshū*. 2 vols. Tokyō: Kazama Shobō.

Fukui, Naoki. 1986. *A theory of category projection and its applications*. Cambridge, MA: MIT dissertation.

Greenberg, Joseph H. 1966. *Language universals: With special reference to feature hierarchies*. The Hague: Mouton.

Hashimoto, Shinkichi. 1934. *Kokugohō yōsetsu* [An outline of Japanese grammar]. Tokyo: Meiji Shoin.

Hashimoto, Shinkichi. 1948. *Kokugohō kenkyū* [Studies on Japanese grammar]. Tokyo: Iwanami Shoten.

Hashimoto, Shinkichi. 1969. *Joshi Jodōshi no kenkyū* [Studies on particles and auxiliaries]. Tokyo: Iwanami Shoten.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kato, Yasuhiko. 1985. *Negative sentences in Japanese. Sophia Linguistica* 19. Tokyo: Sophia University.

Kishimoto, Hideki. 2005. *Tōgokōzō to bunpō kankei* [Syntactic structures and grammatical relations]. Tokyo: Kurosio Publishers.

Kishimoto, Hideki. 2007. Negative scope and head raising in Japanese. *Lingua* 117. 247–288.

Kishimoto, Hideki. 2008. On the variability of negative scope in Japanese. *Journal of Linguistics* 44. 379–435.

Kishimoto, Hideki. 2009. Hobun o toru dōshi to keiyōshi: Kontorōru to jōshō [Complement taking verbs and adjectives: Control and raising]. In Taro Kageyama (ed.), *Nichi ei taishō keiyōshi/fukushi no imi to kōbun* [The meanings of adjectives and adverbs and their constructions], 152–190. Tokyo: Taishukan Shoten.

Kishimoto, Hideki. 2013. Verbal complex formation and negation in Japanese. *Lingua* 135. 132–154.

Kitahara, Yasuo. 1981. *Nihongo jodōshi no kenkyū* [A study of Japanese auxiliaries]. Tokyo: Taishukan Shoten.

Konno, Hiroaki. 2012. The Japanese adjectival conjugational ending drop construction: From a syntax-semantics interface perspective. *Gengo Kenkyu* 141. 5–32.

Konoshima, Masatoshi. 1973. *Kokugo jodōshi no kenkyū* [A study of Japanese auxiliaries]. Tokyo: Ōhūsha.

Kudō, Mayumi. 2004. *Nihongo no asupekuto tensu mūdo taikei: Hyōjungo kenkyū o koe te* [The tense-mood-aspect system in Japanese: Beyond standard Japanese studies]. Tokyo: Hituzi Syobo.

Kuno, Susumu. 1973. *The structure of the Japanese language.* Cambridge, MA: MIT Press.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar, volume 1: Theoretical prerequisites.* Stanford: Stanford University Press.

Martin, Samuel E. 1975. *A reference grammar of Japanese.* New Haven: Yale University Press.

Matsushita, Daizaburō. 1924. *Hyōjun nihon bunpō* [Standard Japanese grammar]. Tokyo: Kigensha.

Matsushita, Daizaburō. 1930. *Kaisen hyōjun nihon bunpō* [Revised standard Japanese grammar]. Tokyo: Tyūbunkan.

Mikami, Akira. 1953. *Gendai gohō josetsu* [Prolegomena to modern usage]. Tokyo: Tōkō Shoin.

Mizutani, Shizuo. 1951. Keiyōdōshi-ben [An account of adjectival nouns]. *Kokugo to Kokubungaku* [Japanese Language and Literature] 28(5). 31–47.

Miyagawa, Shigeru. 1987. Lexical categories in Japanese. *Lingua* 73. 29–51.

Monbushō. 1947. *Chūtōbunpō kōgo* [Intermediate grammar: colloquial language] Tokyo: Chūtōgakkō Kyōkasho.

Muraki, Shinjirō. 2012. *Nihongo no hinshi taikei to sono shūhen* [Japanese parts-of-speech and their surroundings]. Tokyo: Hituzi Syobo.

Ohkado, Masayuki. 1991. On the status of adjectival nouns in Japanese. *Lingua* 83. 67–82.

Okutsu, Kei-ichiro. 1978. *Boku wa unagi da no bunpō* [The grammar of "I'm an eel"]. Tokyo: Kurosio Publishers.

Ōtsuki, Fumihiko. 1889. Gohō shinan [Instructions for usage]. In *Genkai* [Word Sea] (Dictionary of Japanese).

Rodriguez, João. 1604–1608. *Arte da lingoa de Japam.* Translated as *Nihon dai bunten* [A grand book of Japanese grammar] by Tadao Doi, 1955. Tokyo: Sanseido.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch and Barbara L. Lloyd (eds.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Erlbaum.

Sakuma, Kanae. 1951. *Asahi shin kōza: Nihongogaku* [Asahi new series: Japanese linguistics]. Tokyo: Asahi Shinbunsha.

Searle, John. 1969. *Speech acts: An essay in the philosophy of language.* Cambridge: Cambridge University Press.

Shibatani, Masayoshi. 1990. *The languages of Japan.* Cambridge: Cambridge University Press.

Takezawa, Koichi. this volume. Chapter 13: Inflection. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation.* Berlin: De Gruyter Mouton.

Taylor, John R. 1989. *Linguistic categorization: Prototypes in linguistic theory.* Oxford: Clarendon Press.

Teramura, Hideo. 1982. *Nihongo no shintakusu to imi I* [The syntax and the meaning of Japanese I]. Tokyo: Kurosio Publishers.

Tōjō, Gimon. 1841. *Tama no o kuri wake* [Classification of words]. Osaka.

Tokieda, Motoki. 1950. *Nihon bunpō kōgohen* [Colloquial Japanese grammar]. Tokyo: Iwanami Shoten.

Uehara, Satoshi. 1998. *Syntactic categories in Japanese: A cognitive and typological introduction.* Tokyo: Kurosio Publishers.

Uehara, Satoshi. 2003. A diachronic perspective on prototypicality: The case of nominal adjectives in Japanese. In Hubert, Cuyckens, René Dirven, and John R. Taylor (eds.), *Cognitive approaches to lexical semantics*, 363–391. Berlin: Mouton de Gruyter.

Uehara, Satoshi and Fumiko Kumashiro. 2007. *On'in keitai no mekanizumu: Ninchi on'in keitairon no apurōchi* [The mechanism of morpho-phonology: An approach to cognitive morpho-phonology]. Tokyo: Kenkyusha.

Wetzer, Harrie. 1992. "Nouny" and "verby" adjectivals: A typology of predicative adjectival constructions. In Michel Kefer and Johan van der Auwera (eds.), *Meaning and grammar: Cross linguistic perspectives*, 223–262. Berlin: Mouton de Gruyter.

Whaley, Lindsay J. 1997. *Introduction to typology: The unity and diversity of language*. Thousand Oaks, CA: Sage Publications.

Yamada, Yoshio. 1908. *Nihon bunpōron* [On Japanese grammar]. Tokyo: Hōbunkan.

Yamada, Yoshio. 1936. *Nihon bunpōgaku gairon* [Introduction to Japanese grammar]. Tokyo: Hōbunkan.

Yamaguchi, Akiho. 2004. *Nihongo no ronri* [The logic of Japanese]. Tokyo: Taishukan Shoten.

Yumoto, Yoko. 2005. *Fukugō dōshi hasei dōshi no imi to tōgo* [The meaning and the syntax of compound and derivational verbs]. Tokyo: Hituzi Syobo.

Hideki Kobayashi, Kiyo Yamashita, Taro Kageyama

3 Sino-Japanese words

1 Introduction

As described in Chapter 1 (Kageyama and Saito, this volume), the vocabulary of Japanese is composed of four strata: native, Sino-Japanese (S-J), foreign, and mimetic. This chapter outlines the characteristics of Sino-Japanese words or *kango* (漢語), which have a role equal to or in some cases greater than that of native words in word formation in Japanese. Because the previous research in this area has mostly been done in the descriptive framework of traditional grammar, this chapter will put primary emphasis on an overall description of diverse data, with only a few suggestions towards their theoretical analyses.

Sino-Japanese words are words originally borrowed from Chinese and are generally written in Chinese characters (*kanji*).[1] *Kanji* originally developed as logographs or logograms with each character representing a single word or morpheme. For example, the character 木 alone has the meaning of 'tree' and in Sino-Japanese words takes the readings [mokɯ], as in 木馬 *moku-ba* [wood-horse] 'wooden horse' and 材木 *zai-moku* [material-tree] 'wood', or [bokɯ], as in 木刀 *boku-too* [tree-sword] 'wooden sword' and 巨木 *kyo-boku* [giant-tree] 'very tall tree', where both readings are inherited from Chinese. The [m] and [b] phonemic distinction derives from differences in the Chinese pronunciations (Wu dynasty and Han readings) at the times the characters were borrowed. These [mokɯ] and [bokɯ] are not words in isolation; they must always combine with another S-J component to form a word. Single character S-J items are in general then bound morphemes, but there are also many single character S-J words like 茶 *tya* 'tea', 本 *hon* 'book', and 損 *son* 'loss'. Furthermore, among the Sino-Japanese words there are a great number that were created in Japan, in the Meiji period especially.

It is impossible to discuss Sino-Japanese words without some explanation of the *kanji* themselves. When *kanji* were brought into Japan from China around the third and fourth centuries, Japan had no native writing system and the *kanji* were used to represent the native words corresponding to the Chinese meanings of the characters. For example, the native word for 'tree' is *ki* and the character 木 introduced above was used to write it. The native word *ki* in fact appears as the allomorph (apophonic variant) *ko* in some compounds, but the same character 木 is used in such cases as well. As a result, as shown in Table 1, the same single character may both be read as a Chinese word and as a Japanese word.

---

1 The *kanji* used in Japanese often have different forms than the characters used in China and Taiwan today due to independent revisions within Japan.

Table 1: Two ways of reading Kanji characters

	木 'tree'
Chinese reading	*moku*, *boku* (bound morphemes)
Japanese reading	*ki* (free morpheme)

In traditional Japanese grammar, the Chinese reading is called the *on-yomi* (lit. 'sound reading') and the Japanese reading the *kun-yomi* (lit. 'semantic reading'). *On-yomi* show Sino-Japanese morphemes and *kun-yomi* show native morphemes.

Ever since Yamada (1940) discussed S-J words dividing them by the number of *kanji* into one-character, two-character, three-character, and four-character words, there have been quite a number of analyses and categorizations within Japanese research on Sino-Japanese words by the number of characters in a word (e.g. Nomura (1974, 1975, 1988); Zhu (2011, 2013)). This chapter continues this tradition of categorizing and explaining S-J words by the number of characters they contain, but from a linguistics point of view, the following two cautions are necessary.

First, by simply counting characters, there is a danger of overlooking the fact that even when the number of characters is identical, the way the words are formed may be different. For example, the words given in (1a) and (1b) are all composed of two characters and thus, based on appearance alone, would probably all be classified as "two-character S-J words".

(1)　a.　国立 *koku-ritu* [country-build] 'government-funded', 多大 *ta-dai* [many-big] 'great'

　　b.　高校 *koo-koo* [high-school] 'high school', 特急 *tok-kyuu* [special-rapid] 'limited express'

The examples in (1a) are both composed of one Sino-Japanese morpheme added to another S-J morpheme and match in essence the description of two-morpheme Sino-Japanese words. In contrast, although the examples of (1b) are also apparently composed of two characters, they are constructed via a different process of word formation. That is, *koo-koo* is formed from the four-character word 高等学校 *kootoo-gakkoo* 'secondary school' by combining the first character of 高等 *koo-too* 'higher, advanced' with the second character of 学校 *gak-koo* 'school'. The second example, 特急 *tok-kyuu*, is similarly formed through clipping from the four-character word 特別急行 *toku-betu* 'special' *kyuu-koo* 'express'. Clipping is also found in three-character words like 意味深 *imi-sin* 'with profound meaning' (← 意味深長 *imi* 'meaning' *sintyoo* 'profound').

Another danger in classification by the apparent character count is missing regularities in the internal structure of the words. From the term "three-character word", one might expect there simply to be three S-J morphemes (X, Y, and Z) to be lined up as "X-Y-Z", but in fact, except for coordinate compounds like 日米中 *niti-bei-tyuu* 'Japan-U.S.-China', three or more morphemes never occur as a simple linear collocation. For instance, 法改正 *hoo-kai-sei* and 改正法 *kai-sei-hoo* are both composed of three S-J morphemes, but their internal structures differ as shown in (2).

(2) a.   *hoo* 'act, law' + *kai-sei* [change-right] 'amend' → *[hoo [kai-sei]]*
          'amendment of an act'

     b.   *kai-sei* 'amend' + *hoo* 'act' *[kai-sei [hoo]]* 'amended act'
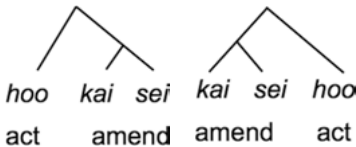
The morphological structures shown in (2a, b) reflect the general rule that complex words are constructed by stacking up "binary branching structures" formed of two elements joined together (cf. Selkirk 1982). Sino-Japanese words composed of four or more characters can fundamentally be analyzed as combinations of pairs of S-J morphemes.

Below, this chapter will examine successively the internal structures and productivity of one-character words (Section 2), two-character words (Section 3), three-character words (Section 4), and four-character words (Section 5). Following that, Section 6 will examine S-J affixes, which are one of the dynamic forces driving word formation in Japanese, and Section 7 will explain clipping in S-J words. Finally, Section 8 will present conclusions and offer some perspectives on future research.

2 Characteristics of one-character words

The main focus of this chapter is on two-, three-, and four-character S-J words, but here we will briefly touch on one-character words.

2.1 Classification of one-character words

One-character S-J morphemes may be free, as in (3a), or bound, as in (3b)–(3f).

(3) a. Free standing nouns: 愛 *ai* 'love', 茶 *tya* 'tea', 本 *hon* 'book', 害 *gai* 'damage', 金 *kin* 'gold', 職 *syoku* 'job', 客 *kyaku* 'customer, guest', 式 *siki* 'ceremony, formula', 会 *kai* 'meeting'

b. Bound connectives: 外出の際 *gaisyutu no* [go.out GEN] **sai** [time, occasion] 'when going out', 車を運転中 *kuruma o* [automobile ACC] *unten* [drive] **tyuu** [while, in the middle] 'when/while driving a car', (朝食) 兼 (昼食) *tyoosyoku* [breakfast] **ken** [doubling as] *tyuusyoku* [lunch] 'both breakfast and lunch at the same time'

c. Morpheme forming the base of a verb: 初物を食す(る) *hatumono o* [first.fruits ACC] **syoku**-*su(ru)* [eat] 'eat the first fruits of the season', 友を信じる *tomo o* [friend ACC] **sin**-*ziru* [trust] 'trust one's friends'

d. Morpheme forming the base of an AN: 急な仕事 **kyuu**-*na* [abrupt-ADN] *sigoto* [job] 'an abrupt job', 別な問題 **betu**-*na* [separate-ADN] *mondai* [problem] 'a separate problem'

e. Morpheme forming the base of an adverb: 実に **zitu**-*ni* [actual-ADV] 'actually', 現に **gen**-*ni* [real-ADV] 'really', 特に **toku**-*ni* [special-ADV] 'especially'

f. Morpheme forming the base of adnominal or adverbial modifiers: 漠とした不安 **baku**-*tosita* [vague-ADN] *huan* [unease] 'a vague unease', 単なる想像 **tan**-*naru* [mere-ADN] *soozoo* [imagination] 'a mere imagination'

Examples in (3a) show single character S-J word standing alone as nouns. The examples in (3b) are of single character S-J words with a conjunctive function; *sai* 'time, occasion' and *tyuu* 'while, in the middle' take a preceding clause with a VN predicate. The single character S-J words in (3c) function as tensed verbs when followed by the native morphemes *su*, *suru*, *zuru*, or *ziru*.

The types exemplified in (3d), (3e), and (3f) are few in number and are unproductive. The examples in (3d) form the stem of an AN (adjectival noun) and take the suffix -*na* to modify nouns. In (3e) the one-character S-J words function as adverbs with -*ni* attached. Accompanied by the classical forms -*tosita* and -*naru*, the words in (3f) are examples of single character S-J words used as adnominal modifiers and have a somewhat formal flavor to them.

2.2 Verb formation with one-character Sino-Japanese words

Consider examples of the type shown in (3c) in which a single character S-J word forms a verb when accompanied by the native suffix -*su(ru)*. At first glance, the

verb 愛する *aisuru* 'love' appears to simply be the one-character S-J word 愛 *ai* 'love' with *suru* 'do' attached. The single character 愛 *ai* functions as an independent noun and *suru* can generally be added to two-character S-J VNs to form verbs, as in 運転する *unten-suru*, so *aisuru* appears to similarly be simply a junction of 愛 *ai* and *suru*. However, Kageyama (1980) showed that there are morphological and syntactic differences between 愛する *aisuru* 'love' and 恋する *koisuru* 'love', which is formed on the native morpheme 恋 *koi* 'love'. First, syntactically, *aisuru* is a transitive verb that takes an accusative object while *koisuru* selects dative case for its object. Morphologically, *koi* and *suru* in *koisuru* can be separated as in (4a), but separating *ai* and *suru* results in ungrammaticality, as shown in (4b).

(4) a. *Takesi*   *wa*   *Naomi*   *ni*   *koi*   *o*   *si-te*   *ir-u.*
         Takeshi   TOP   Naomi   DAT   love   ACC   do-GER   be-PRS
         'Takeshi loves Naomi.'

    b. *Takesi   *wa*   *Naomi*   *ni*   *ai*   *o*   *si-te*   *iru.*
         Takeshi   TOP   Naomi   DAT   love   ACC   do-GER   be-PRS

Koisuru can thus be analyzed in relation to the phrase *koi o suru*, but *aisuru* cannot be made into *ai o suru and must instead be analyzed as a single word.

In fact, comparison of the conjugational forms of *aisuru* and *koisuru* shows this difference clearly; while the negative of *koisuru* is the same as the independent verb *suru*, that of *ai-suru* is different.

(5) a. *suru* 'do'   –   *si-nai* [do-NEG], *sanai

    b. *koisuru*   –   *koisi-nai* [love-NEG], *koisanai

    c. *aisuru*   –   *ais-anai* [love-NEG], *aisinai

Besides *koi-suru*, other verbs that take the pattern shown in (5b) include 損する *son-suru* 'suffer a loss' and 楽する *raku-suru* 'live comfortably'. In contrast, the pattern shown in (5c) with *aisuru* is the same as that shown in the negative forms of native verbs like *kakus-u* 'conceal', *kakus-anai* 'not conceal' (cf. *kakusinai) and *watas-u* 'hand over', *watas-anai* 'not hand over' (cf. *watasinai). Kageyama (1980) concludes from these facts that the stem of *aisuru* is not *ai-* but *ais-*.

Miyake (2010) lists a large number of examples of verbs taking the form "one-character S-J word + *suru*" in their citation forms, some of which are given in (6). (The English glosses give the meanings of the one-character S-J words when used as verbs.)

(6) a. 愛 *ai* 'love', 介 *kai* 'interfere', 期 *ki* 'expect', 害 *gai* 'harm', 帰 *ki* 'attribute', 供 *kyoo* 'serve', 熟 *zyuku* 'ripen', 称 *syoo* 'call', 属 *zoku* 'belong', 題 *dai* 'entitle', 利 *ri* 'benefit', 託 *taku* 'entrust', 要 *yoo* 'need', 有 *yuu* 'have', 適 *teki* 'fit', 類 *rui* 'be similar'

b. 屈 *kuQ* 'give in', 察 *saQ* 'conjecture', 接 *seQ* 'touch', 発 *haQ* 'issue', 熱 *neQ* 'heat', 達 *taQ* 'reach', 反 *haN* 'go against', 面 *meN* 'face'

Based on their conjugations and accent patterns, Miyake (2010) analyzes all of these uniformly as single character S-J words with the verbalizing suffix *-s* attached.

However, examining the morphology of the negative forms closely shows that (6) contains both forms with negatives in *-sanai* and forms with *-sinai*. That is, while all the words in (6a) have negative forms in *-sanai*, all those in (6b) have *-sinai*. The single character S-J words in (6a) end in vowels whereas those of (6b) end in consonants (Q represents the moraic consonant and N the moraic nasal). The generalization can be made that the words in both (6a) and (6b) have in common that they are formed of a single character S-J morpheme with a suffix *-s* attached and their negative forms appear as different allomorphs depending on whether the S-J morpheme ends in a vowel or a consonant.

In this section we examined the lexical categories of one-character S-J words and identified a verbalizing suffix *-s*. This *-s* can turn a single character S-J word into a verb, but it cannot attach to S-J words of two or more characters. As will be shown in the next section, when S-J words of two or more characters are used as verbs, the form *-suru* is used without exception.

3 Characteristics of two-character Sino-Japanese words

This section outlines the characteristics of two-character S-J words composed of two morphemes. Words like 挨拶 *aisatu* for which the meanings of the separate elements 挨 *ai* and 拶 *satu* cannot be specified are excluded. Also not considered are words formed by reduplication like 粛々（と）*syukusyuku(to)* 'solemnly', 満々 *manman (to)* 'brimming with', or 段々（と）*dandan(to)* 'gradually'. Authors like Nomura (1988) working in the framework of traditional Japanese grammar have focussed on the productivity and internal structure of two-character S-J words within the traditional *kokugogaku* approach (Section 3.1). More recent works, such as Kageyama (1980, 1993) and Kobayashi (2004) are based on theoretical morphology and have shifted focus to syntactic aspects such as how two-character S-J words function in a sentence and their ability to take arguments (Sections 3.2 and 3.3).

3.1 Productivity of two-character Sino-Japanese words

When analyzing two-character S-J words, one thing that must be kept in mind is whether the morphemes comprising the two-character S-J word are free or bound. Depending on whether each morpheme is free or bound, four types of words can be identified, as shown in Table 2. In words of Type 1, both morphemes are bound, in Type 2, both are free, and in Type 3 and Type 4 one morpheme is bound and one is free.

Table 2: Composition of two-morpheme S-J words

Type 1. Bound + Bound	*koku ritu* [country-build] 'government-funded', *ka tei* [home-garden] 'family'
Type 2. Free + Free	*gun tai* [army-team] 'armed forces', *zu siki* [picture-formula] 'diagram, chart'
Type 3. Free + Bound	*ai ken* [love-dog] 'one's pet dog', *kin kai* [gold-mass] 'gold ingot'
Type 4. Bound + Free	*zyun ai* [pure-love] 'pure love', *koo tya* [red-tea] '(black) tea'

Of these four types, Type 1 is overwhelmingly the most common in modern Japanese and this section will concentrate on this type.

The productivity of the four types identified in Table 2 varies depending on the individual morphemes comprising the two-character S-J word. For example, the 立 *ritu* 'found, establish' appearing as the second half of 国立 *kokuritu* 'national' can also form 県立 *kenritu* 'prefectural', 私立 *siritu* 'private', 市立 *siritu* 'municipal', 町立 *tyooritu* 'town-owned, precinct-owned', or 村立 *sonritu* 'village-owned'. Similarly, the construction 新-X [new-X] gives 新車 *sin-sya* 'new car', 新人 *sin-zin* 'new person', and 新党 *sin-too* 'new (political) party'. As a prefix, 新 can also appear with S-J words of more than two-characters (e.g. 新世代 *sin-sedai* 'new generation', 新 年生 *sin-itinensei* 'new first-year student'), with native words (新たまねぎ *sin-tamanegi* 'new onions') and with foreign words (新キャベツ *sin-kyabetu* 'fresh cabbage'). Among single character morphemes, then, there are those that contribute to the formation of new words as productive word formation elements. The 激 *geki-* 'intense' that attaches to an action or state to intensify it, besides occurring as a prefix-like element with one-character S-J words like 激走 *geki-soo* 'flat-out run' and 激写 *geki-sya* 'powerful, intense photograph', can also attach to native adjectival stems, as in 激辛 *geki-kara* (*kara-i* 'hot, spicy') and 激うま *geki-uma* (*uma-i* 'yummy').

In general, however, two-character Sino-Japanese words are not very productive. Nomura (1984) compared the 1960 and 1980 editions of a dictionary of neologisms and found only the fourteen examples given in (7) of two-character S-J words that could be considered to have arisen in that 20-year period.

(7) 液晶 *eki-syoo* 'liquid crystal', 学際 *gaku-sai* 'interdisciplinary', 加齢 *ka-rei* 'aging/ageing', 業態 *gyoo-tai* 'business conditions', 性差 *sei-sa* 'sex difference', 声紋 *sei-mon* 'voiceprint', 党友 *too-yuu* 'faction, clique', 劇画 *geki-ga* 'comic strip with dramatic story', 動画 *doo-ga* 'video', 熱波 *nep-pa* 'heat wave', 脳死 *noo-si* 'brain death', 保父 *ho-hu* 'male caregiver in nursery school', 民宿 *min-syuku* 'private home providing lodging for travelers', and 油砂 *yu-sa* 'oil sands'

Many of the words in (7) are created by analogy based on existing two-character S-J words. For example, *gaku-sai* 'interdisciplinary' is based on an analogy with 国際 *koku-sai* 'international', *sei-mon* 'voiceprint' with 指紋 *si-mon* 'fingerprint', and *ho-hu* 'male caregiver' with 保母 *ho-bo* 'female caregiver'. New word creation can be broadly split into rule-governed general "productivity" and more ad hoc "creativity" based on metaphor or analogy (cf. Chapter 1 [Kageyama and Saito, this volume]). Generally speaking, the creation of new two-character S-J words in modern Japanese can probably be viewed as involving analogous creation of new words through operation of a socially influenced speaker's "creativity". Accordingly, when one comes across a new two-character S-J word in the mass media, one is struck by its novelty. In contrast, in the case of four-character S-J words, even an ordinary speaker can create new words comparatively freely. This is a strong indication that formation of four-character S-J words is a productive rule-governed process.

One important standard for the classification of two-character S-J words is whether the entire word can be used independently or not. As shown in Table 2 and the examples of (7), generally two-character S-J words stand as independent words. There are, however, some exceptions to this general rule.

(8) a. 国際 *kokusai*: 国際的 *kokusai-teki* 'international (AN)', 国際性 *kokusai-sei* 'degree of internationalism', 国際関係 *kokusai-kankei* 'international relations'

   b. 積極 *sekkyoku*: 積極的 *sekkyoku-teki* 'positive (AN)', 積極性 *sekkyoku-sei* 'degree of positivity', 積極策 *sekkyoku-saku* 'positive plan'

   c. 消極 *syookyoku*: 消極的 *syookyoku-teki* 'passive, unmotivated (AN)', 消極性 *syookyoku-sei* 'passivity', 消極派 *syookyoku-ha* 'passive/negative faction'

   d. 本格 *honkaku*: 本格的 *honkaku-teki* 'genuine, full-fledged (AN)', 本格派 *honkaku-ha* 'classical school or style', 本格小説 *honkaku-syoosetu* 'serious novel'

Although the examples in (8) are composed of two morphemes, they cannot stand as independent words. However, it is not sufficient to treat these simply as bound

bases, either. In order for these to be used, it is necessary to add a suffix or a noun following them, as shown in (8). Adding a prefix or a noun before them, such as *反国際 *han-kokusai* 'anti-international', *非積極 *hi-sekkyoku* 'non-positive', or *小説本格 *syoosetu-honkaku* 'novel full-fledged', will not qualify them to stand as words. It is not clear how such peculiar morphemes should be handled theoretically.

3.2 Lexical categories of two-character Sino-Japanese words

Not only can most two-character S-J words stand alone as syntactic constituents as shown in (9a), they can also be constituents in the formation of three and four-character S-J words, as shown in (9b).

(9) a. 重要　　　　な　　の　　は　　家庭　で　学習する　　こと　だ
         zyuuyoo　*na*　*no*　*wa,*　***katei***　*de*　***gakusyuu**-suru*　*koto*　*da.*
         important ADN thing TOP home at study-do   thing COP
         'What is important is to study at home.'

     b. 最重要 *sai-zyuuyoo* 'most important', 重要性 *zyuuyoo-sei* 'importance',
         家庭環境 *katei-kankyoo* 'home environment', 課外学習 *kagai-gakusyuu*
         'extracurricular study'

The sentence in (9a) includes three two-character S-J words *zyuuyoo*, *katei*, and *gakusyuu*, all of different lexical categories. *Zyuuyoo* is an adjectival noun (AN) and takes the ending *-na* in its adnominal form and *da* clause-finally. *Gakusyuu* is a verbal noun (VN) and functions as a tensed verb accompanied by the "light verb" *suru*. The categories that have the most members are N, AN, and VN – categories with noun-like properties. Nouns are related to all kinds of concrete objects and abstract notions ranging from human nouns (e.g. 先生 *sensei* 'teacher', 記者 *kisya* 'reporter', 祖父 *sohu* 'grandfather') and artifacts (e.g. 鉛筆 *enpitu* 'pencil', 愛車 *aisya* 'favorite car', 電気 *denki* 'electricity') to natural objects and natural phenomena (e.g. 宇宙 *utyuu* 'universe', 地球 *tikyuu* 'Earth', 地震 *zisin* 'earthquake', 洪水 *koozui* 'flood') to abstract notions (e.g. 哲学 *tetugaku* 'philosophy', 政治 *seizi* 'politics', 平和 *heiwa* 'peace'). ANs include such words as 綺麗 *kirei* 'pretty', 裕福 *yuuhuku* 'wealthy', 稀少 *kisyoo* 'rare', 高度 *koodo* 'high grade', 率直 *sottyoku* 'openhearted', 優秀 *yuusyuu* 'excellent', and 難解 *nankai* 'difficult to understand' that describe attributes, states, or situations and often belong to a comparatively formal register of language. VNs cover a wide range from words of daily use to specialized vocabulary referring to human activities and natural events: 購入 *koonyuu* 'purchase', 研究 *kenkyuu* 'research, study', 自殺 *zisatu* 'suicide', 結婚 *kekkon* 'marry', 失望 *situboo* 'lose hope', 存在 *sonzai* 'exist', 落下 *rakka* 'drop, fall', 誕生 *tanzyoo* 'be born', and many others. In addition, two-character S-J words include

adverbs like 断固 *danko* 'determinedly' and 結局 *kekkyoku* 'ultimately' and interjections like 万歳 *banzai* 'hurray'.

There are a great many two-character Sino-Japanese homonyms. For example, all of 記者 'reporter (N)', 汽車 'steam train (N)', 貴社 'your firm (N)', and 帰社 'return to one's company offices (VN)' are pronounced [kiʃa]. Similarly, 稀少 'rare (AN)', 起床 'arise (VN)', 気象 'weather (N)', and 記章 'badge (N)' are all pronounced [kiʃoː]. In addition, a single two-character S-J word may belong to more than one lexical category, as shown by 迷惑 *meiwaku* in (10).

(10) **Meiwaku**-*na*   *rinzin*   *ga*   *dasu*   *soo'on* *ni*
annoying-ADN neighbor NOM emit noise   DAT

watasi *wa*   **meiwaku**   *si-te*   *iru.*
I   TOP   be.inconvenienced do-GER PRS
'I am inconvenienced by the noise my annoying neighbor makes.'

As apparent by the fact that it has the *-na* ending affixed to it, the first *meiwaku* is an AN and, since it has affixed to it the present progressive of the light verb *suru*, the second *meiwaku* is functioning as a VN. Words that function both as AN and as VN include, besides Sino-Japanese words like 乱暴 *ranboo* 'rough, violent (AN), assault (VN)' and 満足 *manzoku* 'satisfied (AN), satisfy (VN)', words from the mimetic stratum like *unzari* 'tedious (AN), be fed up with (VN)' and *gakkari* 'disappointing (AN), feel disappointed (VN)'.

As described in Chapters 1 (Kageyama and Saito, this volume) and 2 (Kishimoto and Uehara, this volume), verbal nouns (VN) are a hybrid category with characteristics of both nouns and verbs. The verbal character of the VN *gakusyuu* 'learn', for example, is apparent in (11a), in which it co-occurs with *suru* and takes the accusative case marked noun *eigo o* [English ACC] as an object. On the other hand, since *gakusyuu-sya* 'learner' is a noun with the suffix *-sya* affixed to *gakusyuu*, it cannot take an accusative case marked object and the object argument must be realized in the genitive.

(11) a. *eigo*   *o*   *gakusyuu-suru* （学習する）
        English ACC   learn-do
        'to study English'

    b. *eigo*   {*no/*o*}   *gakusyuusya* （学習者）
        English {GEN/*ACC} learner
        'learner of English'

VNs are divided into transitive verbs and intransitive verbs. Table 3 shows VNs subcategorized by the case patterns with which they occur.

Table 3: Classification of VNs according to their case patterns

	Case patterns (*ga*: NOM, *o*: ACC, *ni*: DAT)
Controllable (unergative) intransitive	NP *ga*: 運動 *undoo* 'exercise', 起立 *kiritu* 'stand up', 散歩 *sanpo* 'walk, stroll', 労働 *roodoo* 'labor', 活躍 *katuyaku* 'be active', 激怒 *gekido* 'rage', 動揺 *dooyoo* 'become agitated'
Uncontrollable (unaccusative) intransitive	NP *ga*: 死亡 *siboo* 'die', 蒸発 *zyoohatu* 'evaporate', 発生 *hassei* 'occur, arise', 消滅 *syoometu* 'become extinct, die out', 繁栄 *han'ei* 'flourish', 爆発 *bakuhatu* 'explode', 崩壊 *hookai* 'collapse', 炎上 *enzyoo* 'burst into flames'
Semi-intransitive	NP *ga* NP *ni* or NP *ni* NP *ga*: 到着 *tootyaku* 'arrive', 誕生 *tanzyoo* 'be born', 勤務 *kinmu* 'be employed', 通学 *tuugaku* 'commute to school'
Transitive	NP *ga* NP *o*: 建設 *kensetu* 'build', 運転 *unten* 'drive', 印刷 *insatu* 'print', 調査 *tyoosa* 'survey, investigate', 含有 *gan'yuu* 'include', 購入 *koonyuu* 'purchase', 経験 *keiken* 'experience'
Ditransitive	NP *ga* NP *ni* NP *o*: 提出 *teisyutu* 'submit', 送金 *sookin* 'make monetary transfer', 追加 *tuika* 'append', 郵送 *yuusoo* 'send by mail', 推薦 *suisen* 'recommend', 輸出 *yusyutu* 'export'

3.3 The semantic head of two-character Sino-Japanese words

Finally, let us examine the relation between the two morphemes of two-character S-J words. For convenience' sake, we will employ the term "semantic head" or "semantic core" to refer to the central element that determines the meaning of a complex S-J word (Kobayashi 2004). This diverges from the standard practice in morphology, where the term "head" is defined as the element determining the category of a whole entity containing it (see Chapter 6 [Namiki and Kageyama, this volume]). In the case of two-character S-J words, however, no formal clue is available to identifying the category of the component morphemes, so we have to rely on the semantic relations between the two morphemes in a given S-J word.

For example, the word 愛車 *ai-sya* [love-car] 'one's favorite car' is a noun that refers to a car, not a noun referring to a love. Accordingly, the semantic head of *ai-sya* is the right-hand morpheme *sya* 'car'. In the preponderance of two-character S-J nouns, the semantic head is the right-hand morpheme. VNs and ANs are similar: in the VN 密売 *mitubai* [secretly-sell] 'sell secretly/illegally' (≈密かに売る *hisoka-ni uru* 'sell in secret') and in the AN 病弱 *byoozyaku* [sick-weak] 'having a weak constitution' (≈病気がちで体が弱い *byookigati de karada ga yowai* 'prone to illness with a weak body'), it is the right-hand morpheme that is the semantic core. In comparison, in a coordinate compound (also called "dvandva") like 男女 *danzyo*

[man-woman], since the meaning is 'men and women' neither the left morpheme *dan* 'man' nor the right morpheme *zyo* 'woman' alone can determine the meaning. Accordingly, in the case of coordinate compounds both of the two morphemes comprising the compound can be taken as semantic heads. Such a coordinate relation is not limited to nouns but can be found in other categories as well. Examples include VNs like 救助 *kyuuzyo*, in which both *kyuu* and *zyo* mean 'save', and ANs like 閑静 *kansei* 'quiet', where *kan* and *sei* both mean 'quiet'. The last possibility is that the left morpheme could be the semantic head. This pattern is frequently found in verbal nouns having the structure "verb-like morpheme + noun-like element". For example, since 洗車 *sensya* 'wash-car' does not refer to a type of car but means the act of washing a car, the left-hand morpheme *sen* 'wash' can be taken to be the semantic head. In addition, if one interprets 強力 *kyoo-ryoku* [strong-power] 'powerful' as 力が強い *tikara ga tuyoi* 'power is strong' or 過大 *ka-dai* [over-big] 'too much' as 大き過ぎる *ookisugiru* 'too big', it is probably possible to interpret such ANs as having their semantic heads on the left.

Classifying two-character Sino-Japanese words according to which of the word's two morphemes is the semantic core of the word as a whole yields Table 4.

Table 4: Semantic heads in two-morpheme Sino-Japanese words

	Noun	Verbal Noun	Adjectival Noun
Right-hand semantic head	愛車 *ai sya* [love-car] 'favorite car', 名犬 *mei ken* [name-dog] 'famous dog', 著者 *tyo sya* [write-person] 'author'	密売 *mitu bai* [secret-sell] 'smuggle', 毒殺 *doku satu* [poison-kill] 'kill by poisoning', 骨折 *kos setu* [bone-break] 'break a bone'	病弱 *byoo zyaku* [sick-weak] 'having a weak constitution'
Dual semantic head	男女 *dan zyo* [man-woman] 'men and/or women', 父母 *hu bo* [father-mother] 'parents', 公私 *koo si* [public-private] 'official and personal'	救助 *kyuu zyo* [save-save] 'save', 開閉 *kai hei* [open-close] 'opening and closing'	閑静 *kan sei* [quiet-quiet] 'quiet'
Left-hand semantic head	—	洗車 *sen sya* [wash-car] 'washing a car', 読書 *doku syo* [read-book] 'reading'	強力 *kyoo ryoku* [strong-power] 'powerful'

There are only a few words that might fill the empty position in Table 4 of nouns with left-hand semantic heads: the following could possibly be considered as such.

(12)   *Taroo*   *ga*   ***satu-zin*** (殺人)   *o*   *okasi-ta.*
      Taro   NOM   kill-person   ACC   commit-PST
      'Taro committed murder.'

Since the *satu-zin* [kill-person] in (12) cannot have *suru* attached, that is *satuzin-suru is unacceptable, it is a noun rather than a VN. And, since the meaning of *satu-zin* is not 'a person who kills (or has killed)' but is 'the act of killing a person', the *satu* 'kill' on the left can be considered to be the semantic head.

     Restricting consideration to two-character Sino-Japanese verbal nouns, let us consider their internal structure in more detail. According to Kobayashi (2004), they are roughly categorized as in (13).

(13)   a.   Composed of a verb-like element and a noun corresponding to one of its arguments:
            停電(する) *tei-den(suru)* [stop-electricity (do)] 'blackout, have a power outage', 観戦(する) *kan-sen(suru)* [view-match (do)] 'watch a sports match', 登山(する) *to-zan(suru)* [climb-mountain (do)] 'mountain climb', 読書(する) *doku-syo(suru)* [read-book (do)] 'read books'

      b.   Composed of a verb-like element and another verb-like element:
            救助(する) *kyuu-zyo(suru)* [save-save (do)] 'save', 開閉(する) *kai-hei(suru)* [open-close (do)] 'open and close', 殴殺(する) *oo-satu(suru)* [strike-kill (do)] 'beat to death', 打倒(する) *da-too(suru)* [hit-make.fall (do)] 'knock down, overthrow'

      c.   Composed of an adjunct-like element and a verb-like element:
            病死(する) *byoo-si(suru)* [illness-die (do)] 'die of an illness', 銃殺(する) *zyuu-satu(suru)* [gun-kill (do)] 'shoot to death', 密売(する) *mitu-bai(suru)* [secretly-sell (do)] 'sell illegally', 秒殺(する) *byoo-satu(suru)* [second-kill (do)] 'kill instantly'

Tei-den in (13a) means 電気が停まる *denki ga tomaru* 'electricity stops', where *den* 'electricity' and *tei* 'stop' are in a subject-predicate relation. *Kan-sen* in (13b) means 戦いを観る *tatakai o miru* 'watch a game/match', where *kan* 'watch' and *sen* 'game/ match' are in a predicate-object relation. *To-zan* means 山に登る *yama ni noboru* 'climb onto a mountain', where *to* 'climb' and *zan* 'mountain' are in a predicate-endpoint relation. The relations between the verb-like element and the noun-like element may vary, but in all cases the left-hand verb-like element is the semantic core and the right-hand noun-like element corresponds to an internal argument of the predicate (subject of an unaccusative verb, endpoint of a motion verb, object of a transitive verb).

When two-character Sino-Japanese words of the type shown in (13a) are used as predicates, it may or may not be possible for a noun phrase corresponding semantically to the incorporated argument to appear, as observed by Kageyama (1980, 1999). With *kan-sen* or *to-zan*, the noun-like element contained in the VN can be realized in the sentence as a more concrete referential noun.

(14) a. *Watasi   wa   Hansin-Kyozin-***sen***   o   kan-***sen*** (観戦)
        I   TOP Hanshin.Tigers – Yomiuri.Giants-game ACC spectate
        si-ta.
        do-PST
        'I watched the Tigers-Giants game.'

   b. *Yuuzin ga   Huzi-***san***   ni   to-***zan*** (登山)   *si-ta.*
        friend NOM Fuji-mountain DAT climb-mountain do-PST
        'A friend climbed Mt. Fuji.'

In (14a) the *Hansin-Kyozin-sen* ' Tigers-Giants game' corresponding to the *sen* 'match, game' of *kan-sen* [watch-game] appears in the sentence as an accusative-marked object. Similarly, *Huzi-san* 'Mt. Fuji' corresponds to the *zan* 'mountain' of *to-zan* [climb-mountain] and appears in the sentence as a dative complement. This phenomenon can be called "doubling" and, where the noun-like element internal to the two-character S-J word denotes a general notion ('game' or 'mountain'), the noun phrase appearing in the sentence ('Tigers-Giants game' or 'Mt. Fuji') designates a hyponym to it. This is the most common pattern for two-character S-J VNs used as predicates. There are also cases in which the noun-like element internal to the two-character S-J word and the syntactic object NP stand in a whole-part relation.

(15) a. *Watasi   wa   migi-ude   o   kos-setu* (骨折) *si-ta.*
        I   TOP right-arm ACC break-bone   do-PST
        'I broke my right arm.'

   b. *Titi   wa   tegami   o   kai-huu* (開封)   *si-ta.*
        father TOP letter   ACC open-seal.of.an.envelope do-PST
        'Father opened (the envelope of) a letter.'

In (15a) the object 'right arm' and the 'bone' internal to the compound are in a whole-part relation. (Namely, what actually breaks is not the whole right arm but the 'bone in the right arm' that forms a part of it.) In (15b), 'letter' and 'seal of an envelope' are in the same kind of relation. There are also VNs that do not allow doubling. For example 読書 *doku-syo* [read-book] does not allow an accompanying direct object as in *小説を読書する *syoosetu o doku-syo suru* [novel ACC read-book do] 'read a novel'. Similarly, such collocations as *全力を尽力する *zenryoku o zinryoku suru*

[all.one's.power ACC expend.one's.power do] 'totally expend all one's power' and *汚れた顔を洗顔する *yogoreta kao o sengan suru [dirty face ACC wash-face do] 'wash one's dirty face' are impossible (see also Shimamura 1985).

In the two-character Sino-Japanese words of (13b), which are composed of two verb-like elements, 救助 *kyuu-zyo* 'save' contains two near-synonyms, *kyuu* 'save' and *zyo* 'save', and 開閉 *kai-hei* 'open and close' contains two antonyms, *kai* 'open' and *hei* 'close'. In neither *kyuu-zyo* nor *kai-hei* is it possible to say that one or the other of the verb-like elements determines the meaning of the word as a whole, so, from the perspective of identifying the semantic head, they can probably be analyzed as being dual-head two-character S-J words. In contrast to these, since 殴殺 *oo-satu* [beat-kill] means 'kill by beating', it is the right-hand verb-like element (*satu* 'kill') that can be identified as the semantic head of the word.

Finally, in the examples given in (13c) of two-character Sino-Japanese VNs composed of an adjunct-like element and a verb-like element, since 病死 *byoo-si* [illness-die] means 'to die of an illness', the left-hand adjunct-like element *byoo* can be taken as an adverb-like element giving the cause of death. 銃殺 *zyuu-satu* [gun-kill] means 'kill with a gun (kill by shooting with a gun)' and the left-hand adjunct-like element *zyuu* 'gun' shows the instrument used in the act of killing. 密売 *mitu-bai* [secretly-sell] means 'sell secretly/illegally' and the left-hand adjunct-like element *mitu* shows the manner in which the act of selling takes place. All of these two-character S-J words have their semantic heads on the right.

4 Characteristics of three-character Sino-Japanese words

This section examines the structure and meaning of three-character S-J words. However, as in the case of two-character S-J words, apparent three-character words resulting from clipping of longer expressions, such as 女子大 *zyosidai* (←女子大学 *zyosi-daigaku* 'women's college') and 知事選 *tizisen* (←知事選挙 *tizi-senkyo* 'governor-election') are excluded from consideration.

4.1 Classification of three-character S-J words

The overwhelming majority of three-character S-J words have either the structure "two-character S-J word + one-character S-J word" or the structure "one-character S-J word + two-character S-J word", due to the binary branching structure illustrated in (2) in the introduction to this chapter. Table 5 shows the categories of three-character S-J words having the one-character S-J component on the right (I) or left (II) further sub-categorized by whether or not the one-character S-J morpheme is free (A) or bound (B).

Table 5: Classification of three morpheme S-J words

	A. Free one-character S-J morpheme	B. Bound one-character S-J morpheme
I. One-character S-J morpheme on right	中国 茶 *tyuugoku* **tya** [China-tea], 修正 案 *syuusei* **an** [revision-proposal]	問題 視 *mondai* **si** [problem-view], 永住 権 *eizyuu* **ken** [permanent.residency-right]
II. One-character S-J morpheme on left	腸 閉塞 **tyoo** *heisoku* [intestine-blockage], 市 役所 **si** *yakusyo* [city-offices]	最 先端 **sai** *sentan* [most-leading.edge], 重 労働 **zyuu** *roodoo* [heavy-labor]

Considering the four types given in Table 5 from a semantic perspective, all have their semantic heads on the right. However, they differ in terms of their morphological structure. For example, the word 中国–茶 *tyuugoku-tya* 'Chinese tea' in IA is composed of *tyuugoku* 'China' and *tya* 'tea', and both are free morphemes that can appear as either the left or the right member of a compound (茶–道具 *tya-doogu* 'tea utensils' is an example of a compound with *tya* on the left and 古代–中国 *kodai-tyuugoku* 'ancient China' is an example with *tyuugoku* on the right). Likewise, the words in IIA are composed of free morphemes, so they are probably best analyzed as compounds. On the other hand, the example 問題–視 *mondai-si* 'regard as a problem' in IB can be separated into *mondai* 'problem' and *si* 'view, see as', but *si* is a bound morpheme and moreover, in the meaning of 'view as' can only appear as the right member of a compound, as in 特別–視 *tokubetu-si* [special-view.as] 'regard as special' or 例外–視 *reigai-si* [exception-view.as] 'regard as an exception'. Thus, this 視 *si* in the meaning of 'view, regard as' can be analyzed as a suffix.[2] The word 最–先端 *sai-sentan* in IIB is composed of 最 *sai* 'most' and 先端 *sentan* 'leading edge' and, since *sai* is a bound morpheme that can only appear at the beginning of a word, it can be analyzed as a prefix. Overall, the words in IA and IIA can be analyzed as compounds, and those in IB and IIB as affixed words. Since a single S-J morpheme may have several different uses, however, the distinction between compounding and derivation is often not very clear.

One type of compound in which multiple morphemes co-occur with equal status is coordinate compounds. The examples in (16) are all composed of three Sino-Japanese morphemes that are placed in a coordination relation, but this pattern is not limited to three-character S-J words; there are also cases of two-character S-J words like 男女 *dan-zyo* [man-woman] as seen in Section 3.3 and four-character

---

2 The 視 *si* appearing in 視界 *si kai* [see limit] 'visibility' or 視野 *si ya* [see field] 'field of vision', although written with the same *kanji*, has the meaning 'see, perceive visually' and is not the same as the *si* appearing in *mondai si*.

S-J words like 都道府県 *to-doo-hu-ken* 'the 47 prefectures (from *tookyoo-**to**, hokkai**doo**, oosaka-**hu**, kyooto-**hu**, and 43-**ken**)].

(16) 松竹梅 *syoo-tiku-bai* [pine-bamboo-plum] 'high, middle, and low rankings'
雪月花 *setu-getu-ka* [snow-moon-flowers] 'the beauty of the four seasons'
市町村 *si-tyoo-son* [city-town-village] 'municipalities'

Kageyama (1982) points out a phenomenon notable in Japanese that cannot be properly accounted for within the traditional morphological framework of compounding, derivation, and blending. Examples are given in (17).

(17) a. 与野党 *yo-ya-too* '(political) parties in and out of power' (cf. 与党 *yo-too* 'party in power', 野党 *ya-too* 'party out of power')
内外科 *nai-ge-ka* 'internal medicine and surgery' (cf. 内科 *nai-ka* 'internal medicine', 外科 *ge-ka* 'surgery')
柔剣道 *zyuu-ken-doo* 'judo and kendo' (cf. 柔道 *zyuu-doo* 'judo', 剣道 *ken-doo* 'kendo')
入退院 *nyuu-tai-in* 'entering and leaving hospital' (cf. 入院 *nyuu-in* 'enter the hospital', 退院 *tai-in* 'leave the hospital')
出入国 *syutu-nyuu-koku* 'leaving and entering a country' (cf. 出国 *syuk-koku* 'leave a country', 入国 *nyuu-koku* 'enter a country').

b. 輸出入 *yu-syutu-nyuu* 'export and import' (cf. 輸出 *yu-syutu* 'export', 輸入 *yu-nyuu* 'import')
国内外 *koku-nai-gai* 'domestic and foreign' (cf. 国内 *koku-nai* [country-inside] 'domestic', 国外 *koku-gai* [country-outside] 'foreign')
祖父母 *so-hu-bo* 'grandparents' (cf. 祖父 *so-hu* 'grandfather', 祖母 *so-bo* 'grandmother')

In (17a), 与野党 *yo-ya-too*, for example, is created from *yo-too* 'party in power' and *ya-too* 'party out of power'. Expressed in a more general format, the examples in (17a) are all created from two two-character S-J words *X-A* and *Y-A* that share a common right-hand element *A*, giving a three-character S-J word of the form *X-Y-A*. In (17b), on the other hand, 輸出入 *yu-syutu-nyuu*, for example, is created from *yu-syutu* 'export' and *yu-nyuu* 'import'. The examples in (17b) are all created from two two-character S-J words *X-A* and *X-B* that share a common left-hand element *X*, giving a three-character S-J word of the form *X-A-B*.

How are such examples to be analyzed theoretically? One apparent possibility is the idea of "blending". As shown in Chapter 1 (Kageyama and Saito, this volume), in blending parts of two independent words are extracted and joined to form one word,

as in *gozira* 'Godzilla' from *gorira* 'gorilla' and *kuzira* 'whale'. Just counting the number of morphemes comprising the three-character Sino-Japanese words in (17), the idea of blending may appear promising, but from a semantic perspective this view cannot be adopted. This is because blending normally has the function of creating a new word different from the two original words. While *Godzilla* is a beast completely different from either a gorilla or a whale, *zyuu-ken-doo* 'judo and kendo' does not express a new concept but rather simply expresses the conjunction of the original words *zyuu-doo* 'judo' and *ken-doo* 'kendo'. It is, therefore, not appropriate to analyze *zyuu-ken-doo* as a blending of *zyuu-doo* and *ken-doo*.

Another way of looking at the expressions in (17) would be to analyze them as a combination of a two-character S-J word and a one-character S-J word, that is treat *zyuu-ken-doo* 'judo and kendo' as [*zyuu-ken*] + *doo* or *nyuu-tai-in* as [*nyuu-tai*] + *in*. However, *zyuu-ken and *nyuu-tai do not exist as two-character S-J words. Furthermore, in the case of 内外科 *nai-ge-ka*, the morpheme 外 *ge* 'external' is pronounced not as the more common *gai* but as a special allomorph *ge* that only appears in a very few words like 外科 *geka* 'surgery' and 外宮 *ge-kuu* 'the Outer Shrine of Ise'. 内外 *nai-gai* 'inner and outer' does exist as a two-character S-J word, but if 内外科 were analyzed as 内外 with 科 *ka* 'department' attached, the wrong reading of *nai-gai-ka would be predicted.

Thus, the examples of (17a) are not to be explained with garden-variety word formation processes like simple compounding or blending. There is very little previous research regarding constructions such as these, but Kageyama (1982) analyzes them as coordinate structures incorporating a word-internal empty category. For instance, he argued, by analyzing *nai-ge-ka* as [*nai-Δ_i*] [*ge-ka$_i$*] where the empty category Δ and the *ka* of *ge-ka* are coreferential, the correct interpretation of '*naika* [internal medicine] and *geka* [surgery]' can be obtained. (Note that this coordinate structure incorporating an empty category is the same as the structure proposed for a coordinate structure with predicate deletion, as in *Ani wa ringo o Δ_i, otooto wa banana o tabeta$_i$* 'The older brother Δ_i an apple and the younger brother ate$_i$ a banana'.) A similar structure incorporating an empty category can also be considered for long compounds like those in (18). The symbol ・ (middle dot, centered dot) is a common punctuation mark in Japanese and shows word-internal conjunction (like &). The Δ is part of the analysis and is not realized phonetically or graphically.

(18) a. 格安 Δ・激安航空券 *kakuyasu-*Δ*・gekiyasu-kookuuken* (＝格安航空券 *kakuyasu-kookuuken* [cheap-airline.ticket] & 激安航空券 *gekiyasu-kookuuken* [dirt-cheap-airline.ticket])

   b. 第2 Δ・第3章 *dai-ni-*Δ*・dai-san-syoo* (＝第2章 *dai-ni-syoo* [chapter 2] & 第3章 *dai-san-syoo* [chapter 3])

A problem concerning the analysis of Kageyama (1982) is whether or not hypothesizing a word-internal empty category is theoretically permissible or not. Hypothesizing deletion instead of an empty category gives the same results, but runs into a similar problem of whether or not word-internal deletion is theoretically appropriate. We can conclude, then, that the word formation mechanism observed in (17a) is still unresolved theoretically. In contrast, since the right-hand elements of the examples of (17b) all exist as two-character coordinate compounds, they can be given a simple analysis like (19) without any need to posit empty categories or deletion.

(19) *yu-[syutu-nyuu]* (transport-[out-in]), *koku-[nai-gai]* (country-[in-out]),
       so-[hu-bo] (grand-[father-mother]), *gi-[hu-bo]* (in.law-[father-mother])

4.2 Structure of three-character Sino-Japanese words

Problems regarding the word structure of three-character Sino-Japanese words include the characteristics of each of the two- and one-character S-J words from which they are composed and the ways in which they are combined. Nomura (1974) analyzed three-character S-J words appearing in newspapers and categorized those that were composed of a two-character S-J word followed by a one-character S-J word, such as 問題＋視 *mondai + si* [problem+view.as] or 政治＋家 *seizi + ka* [politics + practitioner] as Type I, which he distinguished from Type II, which consists of a one-character S-J word preceding a two-character S-J word like 最＋先端 *sai + sentan* [most + leading.edge] or 新＋製品 *sin+seihin* [new + product]. Calling the two-character part the "main base" and the one-character part the "sub-base", he further categorized them by the part of speech and the function of the main base as in (20) (parts omitted).

(20) a. Noun: 映画(館) **eiga(kan)** 'movie (theater)', 政治(家) **seizi(ka)** [politics (practitioner)] 'politician'

      b. AN: 貴重(品) **kityoo(hin)** 'valuable (goods)', 有力(者) **yuuryoku(sya)** 'powerful (person)'

      c. VN: 圧倒(的) **attoo(teki)** [overpower (AN-forming suffix)] 'overpowering', 販売(機) **hanbai(ki)** [sell (machine)] 'vending machine'

      d. Adverb: 絶対(的) **zettai(teki)** [absolutely (AN-forming suffix)] 'absolute', 同時(性) **doozi(sei)** 'synchronic(ity)'

      e. Bound morpheme: 顕微(鏡) **kenbi(kyoo)** [microscopic(scope)] 'microscope', 積極(性) **sekkyoku(sei)** 'positive(ness)', 国際(的) **kokusai(teki)** [internation(AN-forming suffix)] 'international', 本格(化) **honkaku(ka)** [serious(change)] 'regularization'

The group of examples in (20e) stand out as being composed of two morphemes, neither of which is free.

Nomura (1974) further divides sub-bases into seven types. Here, we will divide the one-character S-J words corresponding to Nomura's sub-bases first into free and bound morphemes and then by part of speech for a total of nine types. The forms in (21) are all free morphemes that can either be used independently or with a native morpheme affixed.

(21) a. Noun: (宿泊)客 (*syukuhaku*)**kyaku** '(lodging) guest'

     b. Verb: (国内)産 (*kokunai*)**san** '(domestic) production' Cf. 産する *san-suru* 'produce-do'

     c. AN: (問題)別 (*mondai*)**betu** [(problem) separated] 'divided by problem' Cf. 別な *betu-na* 'another'

     d. Adverb: 急(発進) **kyuu**(*hassin*) 'sudden (departure)' Cf. 急に *kyuu-ni* 'suddenly'

The forms given in (22) are all bound morphemes (that is, affixes) and are categorized by their function and use.

(22) a. Noun-forming suffixes: (国際)性 (*kokusai*)**sei** '(internation)ality', (慰霊)祭 (*irei*)**sai** [(comfort spirits)service] 'memorial service', (入学)金 (*nyuugaku*)**kin** [(enter school)money] 'matriculation fee'

     b. AN-forming suffix: (根本)的 (*konpon*)**teki** [(foundation)AN-forming suffix] 'fundamental'

     c. VN-forming suffixes: (絶望)視 (*zetuboo*)**si** [(lose hope)regard.as] 'consider hopeless', (近代)化 (*kindai*)**ka** [(modern times)change] 'modernize'

     d. Prefixes with an adjectival meaning: 低(姿勢) **tei**(*sisei*) 'low (posture)', 新(兵器) **sin**(*heiki*) 'new (weapons)'

     e. Prefixes with a determiner-like meaning: 各(方面) **kaku**(*hoomen*) 'every (direction)', 同(少佐) **doo**(*syoosa*) [same (lieutenant.commander)] 'the same lieutenant commander (mentioned before)'

Care should be taken when classifying to clearly differentiate between those forms that are strictly limited in their distribution to appearing in three-character S-J words and those that are not so limited and can appear in S-J words of any length. Also, it is essential to determine whether the form is limited to appear in S-J words or whether it may also appear with native or foreign words. Table 6 shows that the affixes in (22) can indeed appear with bases other than two-character Sino-Japanese words.

Table 6: Versatility of Sino-Japanese affixes

affix	one-character base	multi-character base	foreign base
-性 *-sei* (noun-forming suffix)	酸性 *san sei* 'acidity'	多民族性 *taminzoku sei* 'multi-ethnicity'	アルカリ性 *arukari sei* 'alkalinity'
-的 *-teki* (AN-forming suffix)	美的 *bi teki* 'aesthetic'	共産主義的 *kyoosansyugi teki* 'communistic'	ヨーロッパ的 *yooroppa teki* 'European'
-視 *-si* 'view as'	敵視 *teki si* 'view as the enemy'	最重要視 *saizyuuyoo si* 'view as most important'	アイドル視 *aidoru si* 'view as pop idol'
-化 *-ka* ('-ize')	液化 *eki ka* 'liquefaction'	軍国主義化 *gunkokusyugi ka* 'shift to militarism'	グローバル化 *guroobaru ka* 'globalization'
新- *sin-* 'new'	新型 *sin gata* 'new pattern'	新千歳空港 *Sin titosekuukoo* 'New Chitose Airport'	新テレビドラマ *sin terebidorama* 'new TV drama'
各- *kaku-* 'each, every'	各局 *kaku kyoku* 'every broadcasting station'	各国立大学 *kaku kokuritudaigaku* 'each national university'	各フライト *kaku huraito* 'every flight'

As demonstrated in Table 6, the affixes in (22) are not limited to appearing in three-character S-J words. One-character Sino-Japanese affixes will be revisited in more detail in Section 6.

The same observation applies to the two-character Sino-Japanese "main bases" given in (20a, b, c, and d). The bases with which these co-occur are not limited to two-character S-J words. Uniquely, the two-character S-J words given in (20e) as bound morphemes only become free-standing words when followed by a one-character suffix or accompanied by a two-character S-J word. 国際 *kokusai* and 本格 *honkaku* co-occur with other types of morphemes comparatively freely, as in 国際シンポジウム *kokusai-sinpoziumu* 'international symposium' (combination with a foreign word) and 本格木造住宅 *honkaku-mokuzoo-zyuutaku* 'genuine wood housing (combination with a four-character S-J word).

5 Characteristics of four-character Sino-Japanese words

Four-character S-J words are words composed of four S-J morphemes, which are typically divided into two words, each consisting of two morphemes. The four

examples shown in (24), 経験豊富 *keiken-hoohu* 'richly experienced', 財務大臣 *zaimu-daizin* 'Finance Minister', 総裁選挙 *soosai-senkyo* 'party chief election', and 出馬表明 *syutuba-hyoomei* 'announcement of candidacy' are of this type.

(24) a. *Keiken-hoohu* (経験豊富) *na* *zaimu-daizin* (財務大臣) *ga*
richly-experienced ADN Finance-Minister NOM

soosai-senkyo (総裁選挙) *ni* *syutuba-hyoomei* (出馬表明) *si-ta*.
party.chief-election DAT candidacy-announcement do-PST
'The highly experienced Finance Minister announced his candidacy for the election for party chief.'

経験豊富 *keikenhoohu* 'richly experienced' is an AN, 財務大臣 *zaimudaizin* 'Finance Minister' and 総裁選挙 *soosaisenkyo* 'party chief election' are nouns, and 出馬表明 *syutubahyoomei* 'announcement of candidacy' is a VN. Each is a phonological unit pronounced with a single accent on the whole. There are also four-character S-J adverbs such as 誠心誠意 *seisin-seii* 'with total sincerity' and 生懸命 *issyoo-kenmei* 'with utmost effort'. This section will explain the structural patterns and productivity of four-character S-J words and will examine their syntactic uses.

5.1 Structural patterns of four-character Sino-Japanese words

It is said that, in theory, Sino-Japanese words of unlimited length can be constructed. However, this limitless productivity is due to repeated application of the compounding patterns "two-character S-J word + one-character S-J word" and "two-character S-J word + two-character S-J word". For example, in the word 新社屋建設案 *sin-syaoku-kensetu-an* 'proposal for construction of new company building', first 新 *sin* 'new' and 社屋 *syaoku* 'company building' combine to make the three-character S-J word *sin-syaoku* 'new company building', which combines with 建設 *kensetu* 'construction' to form the five-character S-J word 新社屋建設 *sin-syaoku-kensetu*. Finally, the six-character S-J word 新社屋建設案 *sin-syaoku-kensetu-an* is formed by adding 案 *an* 'plan, proposal'. The further addition of 発表会 *happyookai* 'presentation' yields the nine-character S-J word 新社屋建設案発表会 *sin-syaoku-kensetu-an-happyookai* 'presentation of plan for construction of new company building', as schematically shown in Figure 1.

The term "four-character Sino-Japanese word" thus should be understood as a shorthand term for "Sino-Japanese words of four or more characters" used for convenience in making comparisons with one-character, two-character, and three-character Sino-Japanese words.

The structures of four-character S-J words can be categorized by the patterns of binary branching structures. Nomura (1975) categorized the structural patterns of
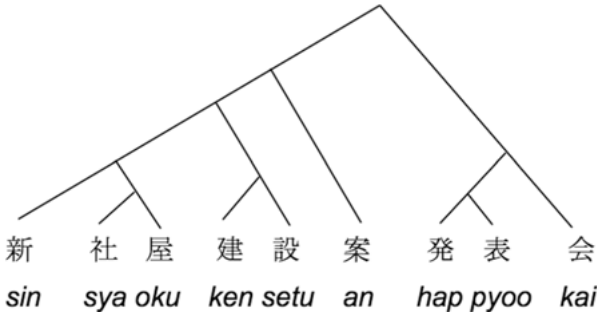
Figure 1: Structure of a nine-character S-J word

four-character S-J using his own notation, which is occasionally difficult to understand. Here his categories have been recast in a more easily comprehensible format using *W*, *X*, *Y*, and *Z* to represent the single character morphemes.

(25) Type I. [[W + X] + [Y + Z]] 上院議員 *zyooingiin* 'member of the Upper House'
[[上 *zyoo* 'upper' + 院 *in* 'chamber'] + [議 *gi* 'deliberate' + 員 *in* 'member']]

Type IIa. [[[W + X] + Y] + Z] 文房具店 *bunbooguten* 'stationery store'
[[[文 *bun* 'letters' + 房 *boo* 'chamber'] + 具 *gu* 'tool'] + 店 *ten* 'store']

Type IIb. [[[W + [X + Y]] + Z] 有資格者 *yuusikakusya* 'person with qualifications'
[[有 *yuu* 'have' + [資 *si* 'asset' + 格 *kaku* 'capacity']] + 者 *sya* 'person']

Type IIc. [W + [[X + Y] + Z]] 新予算案 *sin'yosan'an* 'new budget proposal'
[新 *sin* 'new' + [[予 *yo* 'advance' + 算 *san* 'calculate'] + 案 *an* 'proposal']]

Type IId. [W + [X + [Y + Z]]] 前副首相 *zenhukusyusyoo* 'former vice-premier'
[前 *zen* 'prior' + [副 *huku* 'vice' + [首 *syu* 'chief' + 相 *syoo* 'minister]]]

Type IIIa. [[W・X] + [Y + Z]] 中長距離 *tyuutyookyori* 'mid and long distance'
[[中 *tyuu* 'mid' ・長 *tyoo* 'long'] + [距 *kyo* 'distant' + 離 *ri* 'separate']]

Type IIIb. [[[W・X] + Y] + Z] 農畜産物 *nootikusanbutu* 'farm and livestock products'
[[[農 *noo* 'agriculture'・畜 *tiku* 'livestock'] + 産 *san* 'produce'] + 物 *butu* 'things']

Type IIIc. [[W + X] + [Y・Z]] 巡視船艇 *zyunsisentei* 'patrol ships and boats'
[[巡 *zyun* 'patrol' + 視 *si* 'watch'] + [船 *sen* 'ship'・艇 *tei* 'boat']]

Type IV. [W + X + Y + Z] 都道府県 *todoohuken* 'the 47 prefectures'
[都 *to* 'Tokyo-to' + 道 *doo* 'Hokkaidoo' + 府 *hu* 'Osaka and Kyoto-fu' + 県 *ken* 'all other prefectures']

Type I in (25) is a four-character S-J compound of two two-character S-J words and many four-character S-J words, like those in Table 7 below, are of this type. Type II is composed of a three-character S-J word with a one-character morpheme (prefix or

suffix) attached. For example, since *bunboogu* + *ten* has the same structure as 鮮魚 + 店 *sengyo* + *ten* 'fresh fish + store', it can be treated in a manner comparable to a three-character S-J word; *sin* + *yosan'an* has a structure comparable to the three-character S-J word *sin* + *syaoku*.

Type III structures are all highly unusual four-character Sino-Japanese words. The middle dot symbol (·) used in the formulas for Types IIIa, b, and c shows the meaning "and". *Tyuutyookyori* in Type IIIa can be analyzed as a four-character word resulting from "compressing" 中距離 *tyuukyori* 'middle distance' and 長距離 *tyookyori* 'long distance'. In the case of IIIb, *nootikusanbutu*, 農産 *noosan* 'farm product' and 畜産 *tikusan* 'livestock product' are first compressed giving 農畜産 *nootikusan* 'farm and livestock product' to which 物 *butu* 'things' is added. (Another possibility would be to compress 農産物 *noosanbutu* 'farm products' and 畜産物 *tikusanbutu* 'livestock product'.) Types IIIa and b are examples of the phenomenon described in section 4. Accordingly, they, too, can be analyzed as involving either empty categories or deletion. Type IIIc *zyunsisentei* also involves attaching *sen* 'ship' & *tei* 'boat' after *zyunsi* 'patrol', but, whereas the patterns in Types IIIa and b are very common, it is extremely rare for the head to be a coordinate compound as in Type IIIc. Such words may be restricted to technical terminology. The last type, Type IV, is a coordinate compound with four single character morphemes strung together.

5.2 Categories of four-character S-J words and their internal structure

Four-character Sino-Japanese words corresponding to Type I in (25) nearly all function as nouns, VNs, or ANs. Some examples in common use are given in (26).

(26) a. Nouns
財務大臣 *zaimu-daizin* 'Finance Minister', 鉄道会社 *tetudoo-gaisya* 'railroad company', 家庭環境 *katei-kankyoo* 'home environment', 殺人事件 *satuzin-ziken* 'murder incident', 土地家屋 *toti-kaoku* 'land and buildings'

b. Verbal Nouns
出馬表明 *syutuba-hyoomei* 'declare one's candidacy', 大学改革 *daigaku-kaikaku* 'university reform', 離合集散 *rigoo-syuusan* 'meeting and parting', 受注生産 *zyutyuu-seisan* 'make to order', 意気消沈 *iki-syootin* 'depressed in spirits'

c. Adjectival Nouns
清廉潔白 *seiren-keppaku* 'upright and cleanhanded', 経験豊富 *keiken-hoohu* 'rich in experience', 利用可能 *riyoo-kanoo* 'usable', 単純明快 *tanzyun-meikai* 'simple and clear', 頭脳明晰 *zunoo-meiseki* 'clearheaded'

In the same way as we categorized two-character Sino-Japanese words by the location of their semantic heads in section 3, we can also similarly categorize four-character S-J words. Following Kobayashi (2004), Table 7 shows the categorization of the examples in (26).

Table 7: Semantic heads in four morpheme S-J words

	Nouns	Verbal Nouns	Adjectival Nouns
Right-hand head	財務大臣 *zaimu daizin* 'Finance Minister'	出馬表明 *syutuba hyoomei* 'declare one's candidacy'	経験豊富 *keiken hoohu* 'rich in experience'
Double head	土地家屋 *toti kaoku* 'land and buildings'	通勤通学 *tuukin tuugaku* 'commuting to work and to school'	清廉潔白 *seiren keppaku* 'upright and cleanhanded'
Left-hand head	none	none	none

Regardless of the part of speech, Noun, VN, or AN, the majority of words have right-hand heads. For example, the noun *zaimu-daizin* means 'the minister (*daizin*) in charge of finances (*zaimu*)', the VN *syutuba-hyoomei* means 'announce (*hyoomei*) one's candidacy (*syutuba*)' and the AN *keiken-hoohu(na)* means 'one's experience (*keiken*) is plentiful (*hoohu*)'. Here the right-hand elements *daizin*, *hyoomei*, and *hoohu* are the semantic (and categorical) heads and the left-hand elements function as arguments to or modifiers of the heads. As for the dual-head four-character S-J words, *toti-kaoku* means 'land (*toti*) and buildings (*kaoku*)', *tuukin-tuugaku* means 'commute to work (*tuukin*) and commute to school (*tuugaku*)', and *seiren-keppaku* means 'having integrity (*seiren*) and being clear of blame (*keppaku*)'; in each the two constituents are semantically coordinated.

Among these categories, the verbal nouns behave in a manner closely linked to their syntactic structure. Below we will focus on four-character S-J words functioning as VNs and describe their internal structures and syntactic behaviors, as we did regarding two-character S-J words.

Since the four-character S-J words treated here all have a structure in which one two-character constituent [A-B] is followed by another two-character constituent [X-Y], in describing the internal structure we will refer to the first constituent (two-character S-J word) as [A-B] and the second constituent (two-character S-J word) as [X-Y]. Using this notation, four-character S-J VNs can be divided into three categories as in (27).

(27)　a.　[X-Y] is a VN and [A-B] is a noun representing an argument to the VN.
　　　　地盤沈下(する) *ziban-tinka (suru)* 'land subsidence', 危険回避(する) *kiken-kaihi (suru)* 'danger avoidance', 出馬表明(する) *syutuba-hyoomei (suru)* 'candidacy announcement'

    b.   Both [A-B] and [X-Y] are VNs
       通勤通学(する) *tuukin-tuugaku (suru)* 'commute to work and to school', 冷凍保存(する) *reitoo-hozon (suru)* 'preserve by freezing', 受注生産(する) *zyutyuu-seisan (suru)* 'receive orders and manufacture'

    c.   [X-Y] is a VN and [A-B] is a noun or adjunct
       私費留学(する) *sihi-ryuugaku (suru)* 'self-financed study abroad', 高速回転(する) *koosoku-kaiten (suru)* 'revolve at a high speed', 完全撤退(する) *kanzen-tettai (suru)* 'withdraw completely'

First, in the type in (27a), the first constituents *ziban* 'land' and *kiken* 'danger' are nouns and the second constituents *tinka* 'subside' and *kaihi* 'avoid' are VNs. Unlike the case of two-character S-J VNs, in these four-character S-J VNs the second constituents are the heads. In *ziban-tinka*, meaning *ziban ga tinka suru* 'the land subsides', the unaccusative VN *tinka* 'subside' forms a compound with its subject *ziban* 'land' and in *kiken-kaihi* meaning *kiken o kaihi suru* 'avoid danger' the transitive VN *kaihi* 'avoid' forms a compound with its object *kiken* 'danger'. The subjects of unaccusative verbs and the objects of transitive verbs are called "internal arguments" and in *ziban-tinka* and *kiken-kaihi* the internal arguments have joined with the VNs. In contrast, there is a general constraint that "external arguments" (the subjects of unergative intransitive verbs and the subjects of transitive verbs) do not occur inside compounds and, in fact, in four-character S-J VNs as well, external arguments are excluded (Kageyama 1993; Kobayashi 2004).

(28)   a. *作家自殺(する) **sakka-zisatu (suru)* 'author-suicide' (cf. *Sakka ga zisatu suru* 'an author commits suicide.')

      b. *医師診察(する) *isi-sinsatu 'doctor-examination' (cf. *Isi ga kanzya o sinsatu suru* 'The doctor examines the patient.')

However, transitive VNs compounded with their subjects are observed, apparently violating the above generalization.

(29)   a. 総理主催 (のパーティ) *soori-syusai (no paatii)* '(a party) hosted by the prime minister' (cf. *Soori ga paatii o syusai suru* 'The prime minister hosts a party.'

      b. 学長主導 (の大学改革) *gakutyoo-syudoo (no daigaku-kaikaku)* 'university reform led by the university president' (cf. *Gakutyoo ga kaikaku o syudoo suru*. 'The university president leads the reform.')

Looking only at the relations within the compound, *soori-syusai* and *gakutyoo-syudoo* in (29) appear to be formed of a transitive VN and its subject. Interestingly,

however, as shown in (30), these compounds cannot form a tensed verb through the affixation of *suru*.

(30)   a.   *総理主催したパーティ *soori-syusai sita paatii*
              'the party the prime minister hosted'

        b.   *学長主導した大学改革 *gakutyoo-syudoo sita daigaku-kaikaku*
              'the university reform the university president led'

That is, *soori-syusai* and *gakutyoo-syudoo* are not VNs after all, but, as shown in (29) are a kind of adjective-like noun that carries the function of modifying the following noun (*paatii* 'party' or *daigaku-kakumei* 'university reform'). This type of compound is not limited to four-character S-J words, but is widely used (see Chapter 7 [Kageyama, this volume]).

    The four-character S-J words in (27b) are composed of two VNs. Since *tuukin-tuugaku* means *tuukin* 'commute to work' and *tuugaku* 'commute to school' and can be considered to have a coordinate structure, from the point of view of identifying the semantic head, it can be taken as having a dual-head structure. Cases like 暴飲暴食 *booin-boosyoku*, meaning *booin* 'heavy drinking' and *boosyoku* 'heavy eating', and 連携協力 *renkei-kyooryoku*, meaning *renkei* 'coordinate' and *kyooryoku* 'cooperate', can also be analyzed as having coordinate structures. However, not all four-character S-J VNs composed of two VNs have a coordinate (i.e. dual-head) structure. *Reitoo-hozon* and *zyutyuu-seisan* given in (27b) can be paraphrased as *reitoo site hozon suru* 'preserve by freezing' and *zyutyuu site seisan suru* 'manufacture after receiving orders', respectively, and so the second constituents are clearly the semantic heads.

    Finally, in the four-character S-J words of the type in (27c), the first constituents have either adjunct-like or adverbial meanings and serve to modify the second constituent VNs as demonstrated by the fact that, for example, *sihi-ryuugaku* 'self-financed study abroad' can be restated as *sihi de ryuugaku (suru)* 'study abroad on private funds', *koosoku-kaiten* as *koosoku de kaiten (suru)* 'revolve at a high speed', and *kanzen-tettai* as *kanzen ni tettai (suru)* 'withdraw completely'. Furthermore, the whole four-character S-J word is a VN as are the second constituents. From these facts, it is clear that the second constituents are heads both semantically and categorically.

    To conclude, except for a small number of VNs with a coordinate structure, four-character S-J VNs all have a right-hand head structure.

5.3 Syntactic uses of four-character Sino-Japanese verbal nouns

Like two-character and three-character Sino-Japanese verbal nouns, four-character S-J VNs as a whole function syntactically as transitive or intransitive verbs and can take subjects and objects. With four-character S-J VNs, when the left-hand morpheme is an adjunct (like *zidoo* 'automatic') and the right-hand is a VN (like *unten* 'drive'), as in 自動−運転 *zidoo-unten* 'automated driving', the arguments are naturally deter-

mined by the VN. In contrast, when the four-character S-J VN is composed of two VNs, there is a question as to which determines the arguments. Two possibilities can be thought of. First, using 体験－入部 *taiken-nyuubu* 'trial membership in a club' as an example, the left VN *taiken* 'experience' takes an accusative object when used as an independent VN, as in (31a), while the right VN *nyuubu* 'join a club' takes a dative complement when used independently, as in (31b).

(31) a. *Musuko   ga   zyuudoo-bu   {o/*ni}   taiken   si-ta.*
       son   NOM   judo-club   {ACC/*DAT}   experience   do-PST
       'My son tried the judo club out.'

    b. *Musuko   ga   zyuudoo-bu   {ni/*o}   nyuubu   si-ta.*
       son   NOM   judo-club   {DAT/*ACC}   enter.club   do-PST
       'My son joined the judo club.'

As shown in (32), though, *taiken-nyuubu* 'trial membership in a club' can only take a dative complement.

(32) *Musuko   ga   zyuudoo-bu   {ni/*o}   taiken-nyuubu   si-ta.*
    son   NOM   judo-club   {DAT/*ACC}   trial.membership.in.a.club   do-PST
    'My son joined the judo club on a trial basis.'

Thus, it is the second constituent, *nyuubu* 'join a club', that determines the case marking in (32). Four-character S-J words that have *nyuubu* as a second constituent all function as VNs taking *-suru* 'do' whether or not the first constituent can take *-suru* (is a VN), as shown by a comparison of *taiken-nyuubu (suru)*, the first constituent of which does take *-suru* (*taiken-suru* 'experience') and 時－入部 (する) *itizi* 'temporary' – *nyuubu* 'join a club', the first constituent of which does not (**itizi-suru*). In this type of four-character S-J VN, then, it is the second constituent that is the head of the whole VN and which determines its syntactic category and, in an example like (32), it is the right-hand head *nyuubu* that determines the case marking.

    Fundamentally, it is the head VN that determines argument selection, but there are also cases in which the four-character S-J VN as a whole inherits the arguments of the non-head (i.e. left-hand) VN. For example, *syutuba-hyoomei* 'announce one's candidacy' takes a dative complement, as in (33).

(33) *Zaimu-daizin   ga   soosai-senkyo   {ni/*o}   syutuba-hyoomei*
    Finance-minister   NOM   party.chief-election   {DAT/*ACC}   candidacy-announce
    si-ta.
    do-PST
    'The Minister of Finance announced his candidacy for party president.'

In contrast, *syutuba* and *hyoomei* used as independent VNs take the case marking shown in (34).

(34) a. *Zaimu-daizin   ga   soosai-senkyo   {ni/*o}   syutuba   si-ta.*
   Finance-minister NOM party.chief-election {DAT/*ACC} candidacy do-PST
   'The Minister of Finance became a candidate for party president.'

  b. *Zaimu-daizin   ga   {syutuba/*soosai-senkyo}   o*
   Finance-minister NOM {candidacy/*party.chief-election} ACC

   hyoomei   si-ta.
   announce do-PST
   'The Minister of Finance announced his {candidacy/*the election for party president}.'

That *syutuba* 'become a candidate' takes a dative complement and not an accusative object is shown by (34a). On the other hand, as shown in (34b), *hyoomei* 'announce' takes an accusative object. Since the case marking assigned by the four-character S-J VN *syutuba-hyoomei* as a whole is dative, as shown in (33), this case marking must come not from the head *hyoomei*, but from the non-head *syutuba*. Not only that, but the noun *soosai-senkyo* appearing as the dative argument of the four-character S-J *syutuba-hyoomei* as a whole is identical to the noun taken as an argument by *syutuba* alone in (34a). In a case like that of *syutuba-hyoomei*, since both constituents *syutuba* and *hyoomei* are VNs it is impossible to determine which is the head that determines the category of the four-character S-J word as a whole, but in a case like *syosin-hyoomei (suru)* [opinion-announcement (do)] 'statement of conviction', since an expression like *syosin-suru 'opinion-do' is not possible, it is clear that in four-character S-J words with *hyoomei* as the second constituent, it is the right-hand VN *hyoomei* that functions as the head to determine the syntactic category (VN) for the word as a whole. Thus, it is clear that the argument *soosai-senkyo* in (33) is inherited from the non-head constituent *syutuba*.

  Why is an argument inherited not from the head but from a non-head? What is crucial is that in the four-character S-J word *syutuba-hyoomei*, the non-head *syutuba* corresponds to the object of the head *hyoomei* (cf. *syutuba o hyoomei (suru)* [candidacy ACC announce (do)] 'announce one's candidacy') and the demand of *hyoomei* for an accusative object is satisfied within the four-character S-J word. In contrast to this, since *syutuba* takes arguments as a VN in its own right, as in (34a), when it becomes an internal constituent within a four-character S-J VN, it becomes necessary for its argument to be realized in the sentence. It is in this way that an argument can be inherited from a non-head. Inheriting an argument from a non-head is also observed with native words, as discussed in Chapter 14 (Kageyama, this volume).

6 Sino-Japanese affixes

The concept of "Sino-Japanese affixes" in Japanese was clearly specified as a word-level unit in the section titled "Gohō shinan" (Instruction in grammar) printed in the introduction to ŌTSUKI Fumihiko's *Genkai* (1882), the first dictionary of Japanese. In addition to eight parts of speech, there were listed *settōgo* (prefix words) and *setsubigo* (suffix words) within each of which several Sino-Japanese prefixes and suffixes were listed in addition to native words. This section will consider as S-J affixes primarily single character S-J affixes that attach to two-character S-J words. It is important to note, however, that many S-J affixes are not limited in use to Sino-Japanese words but also attach to native and foreign bases as well. This point was already shown in Table 6, and additional examples are given below in (35).

(35) a. [Pref. + S-J base] 最下位 ***sai**-kai* [most-low.rank] 'the lowest rank'

[Pref. + Foreign base] 全ヨーロッパ ***zen**-yooroppa* [whole-Europe] 'the whole Europe'

[Pref. + Native base] 再貸し出し ***sai**-kasidasi* [again-lend] 'lend out again'

b. [S-J base + Suf.] 改革派 *kaikaku-**ha*** [reform-faction] 'reformist faction'

[Foreign base + Suf.] カナダ人 *Kanada-**zin*** [Canada-person] 'Canadian'

[Native base + Suf.] 田舎風 *inaka-**huu*** [country-style] 'rustic'

This section will first examine the semantic characteristics of S-J affixes in Section 6.1 and then discuss their morphological characteristics in Section 6.2.

6.1 Semantic characteristics of Sino-Japanese affixes

Affixes are categorized into prefixes and suffixes by whether they appear as the left-hand or right-hand constituents of the derived word. However, in the case of one-character S-J words, there are some that can appear in both the left and right positions as shown in (36).

(36) a. 御意見 ***go**-iken* '(honorific) opinion', 姪御 *mei-**go*** '(someone else's) niece'

b. 上半身 ***zyoo**-hansin* 'upper torso', 上出来 ***zyoo**-deki* 'well-made', 立場上 *tatiba-**zyoo*** 'given one's position'

c. 大発見 ***dai**-hakken* 'great discovery', 米粒大 *kometubu-**dai*** 'rice-grain-sized'

The *go* appearing in (36a) is an archetypical example of a Sino-Japanese prefix, but it is also used as a suffix. In both uses it has the common meaning of showing politeness or respect. However, the underlined single character S-J morphemes in (36b, c) each have different meanings. 上 *zyoo* means 'upper' in *zyoo-hansin* and 'good' in *zyoo-deki*. In *tatiba-zyoo*, the meaning of *zyoo* is more abstract, along the lines of 'on the point that …, on the basis of …'. *Dai* in *dai-hakken* means 'big, great' but in *kometubu-dai* it means 'size, bigness'. Opinions are divided as to whether the *zyoo* in (36b) and the *dai* in (36c) should be analyzed as affixes or as compound constituents. There are quite a few such cases concerning S-J words where the distinction between affix and compound constituent is not clear.

It is often thought that affixes, unlike content words, are devoid of substantive meaning. However, since Sino-Japanese morphemes from the beginning carry specific meanings, compared to affixes in English or native affixes, they often show substantive semantic concepts. Meanings carried by affixes are shown in (37) and (38). The examples in (37) are prefixes and those in (38) are suffixes.

(37) a. The prefix limits or modifies the meaning of the base.

　　　　貴(大学) **ki**(*daigaku*) 'your (university)', 好(成績) **koo**(*seiseki*) 'good (results)', 新(発売) **sin**(*hatubai*) 'newly (for sale)', 高(濃度) **koo**(*noodo*) 'high (concentration)', 遠(赤外線) **en**(*sekigaisen*) 'far (infrared)', 低(レベル) **tei**(*reberu*) 'low (level)'

　　 b. The prefix has a verbal meaning and the base noun corresponds to its argument.

　　　　反(体制) **han**(*taisei*) 'anti(establishment)', 助(監督) **zyo**(*kantoku*) 'assistant (director)', 有(酸素) **yuu**(*sanso*) [having (oxygen)] 'aerobic', 脱(サラリーマン) **datu**(*sarariiman*) [escape (salaried worker)] 'setting oneself free from a white-collar job', 省(エネルギー) **syoo**(*enerugii*) 'reduce (energy)'

　　 c. The prefix limits the reference of the base noun ("determiner-like" prefixes).

　　　　本(製品) **hon**(*seihin*) 'this (product)', 同(大学) **doo**(*daigaku*) [same (university)] 'the previously mentioned university', 前(総理大臣) **zen**(*sooridaizin*) 'ex-(prime minister)', 各(参加者) **kaku**(*sankasya*) 'every (participant)',

　　 d. The prefix modifies a predicate-like base adverbially.

　　　　急(成長) **kyuu**(*seityoo*) 'rapid (growth)', 超(過激) **tyoo**(*kageki*) 'super(radical)', 過(保護) **ka**(*hogo*) 'over(protective)', 再(検討) **sai**(*kentoo*) [again(consider)] 'reconsider', 激(辛) **geki**(*kara*) 'intensely (spicy)'

e. The prefix shows negation.

不(確か) ***hu***(*tasika*) 'un(certain)', 無(責任) ***mu***(*sekinin*) 'ir(responsible)',
未(経験) ***mi***(*keiken*) [not.yet(experience)] 'inexperienced', 非(正規雇用)
hi(*seiki-koyoo*) 'non(regular employment)'

Prefixes of the type in (37a) are comparatively common. Prefixes like this function to modify the right-hand constituent. Such prefixes are also found in native words as in 大騒ぎ ***oo***-*sawagi* 'big uproar' and 小雨 ***ko***-*same* 'light rain'. Prefixes of the type in (37b) are not as common as those in (37a). The internal structure of the derived word takes the form of a Chinese [V + N] syntactic constituent formed of the verb-like prefix and the noun corresponding to its argument. The type of prefixes in (37c) are characteristically Sino-Japanese; the only native prefix of this type being the 元 *moto* of 元(会長) ***moto***(*kaityoo*) 'former (chairman)'. The determiner-like prefixes of (37c) also have unique morphological and syntactic characteristics (see Chapter 14 [Kageyama, this volume] for details). The type of S-J prefix in (37d) is also not common and is even more uncommon in native words with ほの(暗い) ***hono***(*gurai*) [faintly(dark)] 'gloomy' being about the only example. The category in (37e) is separate from the concept of part of speech. The prefixes of this category are all comparatively productive and share the meaning of negation. There being no native stratum negative prefixes, these are characteristically Sino-Japanese prefixes and have been placed in a single group. The internal structure of the derived words is the same as Chinese syntactic constituency. Considering all of the above factors, we can conclude that, except for those of the type in (37a), prefixes are uniquely Sino-Japanese.

(38) a. The suffix has a noun-like meaning and derives a noun.

(作曲)家 (*sakkyoku*)***ka*** [(compose) person] 'composer', (図書)館 (*tosyo*)***kan*** [(books) edifice] 'library', (乗車)券 (*zyoosya*)***ken*** '(boarding) ticket', (バス)代 (*basu*)***dai*** '(bus) fare', (アルカリ)性 (*arukari*)***sei*** [(alkaline)ness] 'alkalinity', (山田)氏 (*Yamada*)***si*** [(Yamada) mister] 'Mr. Yamada'

b. The suffix has a verb-like meaning and derives a noun or a VN.

(北海道)産 *(hokkaidoo)****san*** [(Hokkaido) product] 'product of Hokkaido', (中国)製 (*tyuugoku*)***sei*** [(China) manufacture] 'made in China', (収入)減 (*syuunyuu*)***gen*** '(income) reduction', (日光)浴 (*nikkoo*)***yoku*** [(sun light) bathing] 'sunbathing', (疑問)視 (*gimon*)***si*** [(doubt) view] 'view as doubtful', (無人)化 (*muzin*)***ka*** [(uninhabited) change] 'depopulate'

    c.   The suffix adds the concept of "type" or "style" to the base noun deriving an AN or a noun.

(日本)的 (*nihon*)***teki*** [(Japan)like] 'Japanese', (振り子)式 (*huriko*)***siki*** '(pendulum) type', (サラリーマン)風 (*sarariiman*)***huu*** [(white-collar worker) style] 'white-collar style', (文語)調 (*bungo*)***tyoo*** [(literary language register] 'literary-sounding'

    d.   The suffix attaches to a noun and derives an adverb.

(事実)上 (*zizitu*)***zyoo*** [(fact) on] 'based on the facts', (戦時)下 (*senzi*)***ka*** [(wartime) under] 'under wartime conditions', (東京大阪)間 *Tookyoo-Oosaka*)***kan*** [(Tokyo-Osaka) between] 'between Tokyo and Osaka', (時間)外 (*zikan*)***gai*** [(hours) outside] 'outside of regular hours', (紳士)然 (*sinsi*)***zen*** [(gentleman) state] 'gentlemanly'

    e.   Noun classifiers

(馬2)頭 (*uma ni*)***too*** [(horse 2) head] 'two horses', (ビール3)本 (*biiru san*)***bon*** [(beer.bottles 3) long.thing.counter] 'three bottles of beer', (椅子2)脚 (*isu ni*)***kyaku*** [(chairs 2) furniture.with.legs.counter] 'two chairs', (テレビ1)台 (*terebi iti*)***dai*** [(television 1) mechanical.contrivance.counter] 'one television set', (スーツ1)着 (*suutu it*)***tyaku*** [(suit 1) suits.counter] 'one suit'

The suffixes in (38a) form derived nouns and many S-J suffixes are included in this type. The suffixes *-san*, *-sei*, *-gen*, and *-yoku* in (38b) derive nouns and *-si* and *-ka* derive verbal nouns. Those in (38c) derive ANs, but their number is not large. The examples in (38d) are affixes deriving adverbs but many of this type would be better analyzed as clitics than as affixes. Those in (38e) are S-J numeral classifiers. Many of the suffixes in (38) do not add semantic characteristics but can be thought to determine the category of the derived word as a whole.

6.2 Morphological characteristics of Sino-Japanese affixes

Important as characteristics of affixes are their function of changing the category of the base and lexical constraints on the base. Sino-Japanese affixes basically attach to a noun base to form a noun, but there are cases in which parts of speech other than nouns are derived depending on syntactic constituents that follow. There are not a large number of category-changing S-J affixes; the primary ones are given in Table 8.

Table 8: Category-changing S-J affixes

Prefixes that change a noun into an AN	無 *mu* 'not' + 意味 *imi* 'meaning' → 無意味(な) ***mu*** *imi*(*na*) 'meaningless' 不 *hu* 'not' + 規則 *kisoku* 'regular' → 不規則(な) ***hu*** *kisoku*(*na*) 'irregular' 大／小 *dai/syoo* 'large/small' + 規模 *kibo* 'scale' → 大／小規模(な) ***dai/syoo*** *kibo*(*na*) 'large/small scale' 有 *yuu* 'having' + 意義 *igi* 'significance' → 有意義(な) ***yuu*** *igi*(*na*) 'significant'
Suffixes that change a noun into a VN	疑問 *gimon* 'doubt' + 視 *si* 'view as' → 疑問視(する) *gimon* ***si*** (*suru*) 'consider doubtful' 近代 *kindai* 'modern' + 化 *ka* 'change' → 近代化(する) *kindai* ***ka***(*suru*) 'modernize'
Suffix that changes a noun into an AN	道徳 *dootoku* 'morals' + 的 *teki* → 道徳的(な) *dootoku* ***teki***(*na*) 'moral (adj)'
Suffixes that change a noun into an adverb	事実 *zizitu* 'fact' + 上 *zyoo* 'on' → 事実上 *zizitu* ***zyoo*** 'factually' 戦時 *senzi* 'wartime' + 下 *ka* 'under' → 戦時下 *senzi* ***ka*** 'under wartime conditions'

Prefixes that have the function of changing categories are, in addition to the negative prefixes *mu* and *hu*, limited to some that show size or existence, like *dai*, *syoo*, and *yuu*. Compared to prefixes, category-changing suffixes are more numerous, but the proportion of all suffixes they occupy is limited.

Next let us consider restrictions that affixes place on their bases. In general, except for the noun classifiers, S-J affixes attach freely to both single character and two-character S-J bases, as seen in 再会 ***sai***-*kai* 'meet again', 再会談 ***sai***-*kaidan* 'discuss again' and 和式 *wa*-***siki*** 'Japanese style', 西洋式 *seiyoo*-***siki*** 'Western style'. Also, although the bases to which they attach are overwhelmingly Sino-Japanese words, there also many that can attach to native and foreign words as well as S-J ones. However, as there are a wide variety of Sino-Japanese affixes, ranging from those for which the bases they can combine with are severely restricted to those that attach quite freely, with productivity also varying widely, it is necessary to analyze each affix individually. Here, we will first take up the negative prefixes. There have been comparatively many proposals concerning negative prefixes, such as Nomura (1973) and Kageyama (1982, 1999). Table 9 is based on Nomura (1973) and shows the categories of the bases each prefix may combine with. The symbol * indicates a non-occurring combination.

無 *mu*- has the meaning 'lacking, non-existent' and attaches to nouns and verbal nouns but not to adjectival nouns. 未 *mi*-, which means 'an action has still not reached completion', attaches to a few nouns but mostly to VNs. 不, read *hu* or *bu*, shows negation much like English *in*- as in *insignificant* or *un*- as in *unkind* and is the most productive, attaching to nouns, adjectival nouns, and verbal nouns. 非 *hi*- mostly expresses a meaning corresponding to English *non*- and attaches to nouns, ANs, and VNs.

Table 9: Negative prefixes and the categories of their bases

Prefix	N bases	AN bases	VN bases
無 *mu-*	無 関心 **mu kansin** 'unconcerned', 無 神経 **mu sinkei** 'insensitive'	*	無 関係 **mu kankei** 'unrelated', 無 理解 **mu rikai** 'lack of understanding'
未 *mi-*	未 成年 **mi seinen** 'not of age', 未 払い **mi harai** 'unpaid'	*	未 経験 **mi keiken** 'not yet experienced', 未 解決 **mi kaiketu** 'unresolved'
不 *hu-*	不 景気 **hu keiki** 'recession', 不 手際 **hu tegiwa** 'ineptitude'	不 確実 **hu kakuzitu** 'uncertain', 不 平等 **hu byoodoo** 'unequal'	不 承知 **hu syooti** 'disapproval'. 不 勉強 **hu benkyoo** 'lack of study'
非 *hi-*	非 人情 **hi ninzyoo** 'inhuman', 非 常識 **hi zyoosiki** 'lack of common sense'	非 合法 **hi goohoo** 'unlawful', 非 衛生的 **hi eiseiteki** 'unsanitary'	非 公認 **hi koonin** 'unauthorized', 非 軍事化 **hi gunzika** 'non-militarized'

All of the negative prefixes can attach to one-character and two-character Sino-Japanese bases as in 無‐限 **mu**-*gen* 'infinite', 未‐定 **mi**-*tei* 'not yet fixed', 不‐利 **hu**-*ri* 'disadvantage', 非‐力 **hi**-*riki* 'powerless' and 無‐期限 **mu**-*kigen* 'indefinite', 未‐提出 **mi**-*teisyutu* 'not yet submitted', 不‐衛生 **hu**-*eisei* 'unhygienic conditions', 非‐会員 **hi**-*kaiin* 'non-member'. *Hi-* can further attach to three-character S-J words as in 非‐人間性 **hi**-*ningensei* 'inhumanity' and 非‐製造業 **hi**-*seizoogyoo* 'non-manufacturing industry'. In addition, *hi-* is the only one that can attach to Sino-Japanese words of four or more characters or to foreign words (非‐民主主義 **hi**-*minsyusyugi* 'undemocratic', 非‐リアル **hi**-*riaru* 'unreal').

Consider now two suffixes meaning 'person': 人 -*zin* and 人 -*nin*. Nomura (1977) shows that they clearly contrast in terms of word formation, standing in complementary distribution. Namely, whereas the first attaches to both noun and AN stems as in 外国‐人 *gaikoku-zin* 'foreigner', 芸能‐人 *geinoo-zin* 'performer', and 有名‐人 *yuumei-zin* 'famous person', the latter only attaches to VN stems, as in 通行‐人 *tuukoo-nin* 'passerby', 弁護‐人 *bengoo-nin* 'defender', 管理‐人 *kanri-nin* 'manager'. Furthermore, there is a restriction on the type of word that can be a base. 人 -*nin* can also attach to native words, as in 受け取り‐人 *uketori-nin* 'recipient', 立ち会い‐人 *tatiai-nin* 'witness', and 付き添い‐人 *tukisoi-nin* 'attendant'. With the sole exception of 暇‐人 *hima-zin* 'person of leisure', 人 -*zin* never combines with native words.

A representative, highly productive Sino-Japanese affix that combines with a variety of bases is the suffix -的 -*teki*, which makes an adjectival noun. Besides affixing to S-J bases unrestricted by the number of characters in the base, as shown by 私‐的 *si-teki/watasi-teki* 'private/personal', 活動‐的 *katudoo-teki* 'active',

政治家-的 *seizika-teki* 'politician-like', and 共産主義-的 *kyoosansyugi-teki* 'communistic', it can also form ad hoc derivations by combining with native words, foreign words, mixed types, and proper nouns, as in 草分け-的 *kusawake-teki* 'pioneering', カリスマ-的 *karisuma-teki* 'charismatic', 手前味噌-的 *temae-miso-teki* 'self-flattering', and 寅さん-的 *Tora-san-teki* 'Tora-san-like'. Most of the bases *-teki* combines with are nouns or verbal nouns with comparatively few examples of adjectival noun bases. This is because there is no need to convert to an AN a base that already has AN characteristics. Both 健康 *kenkoo* (AN) and 健康的 *kenkoo-teki* (AN) are found, but the meanings they convey are different: *kenkoo-na* means 'healthy, physically free of disease' whereas *kenkoo-teki-na* means 'healthful, aids in promoting bodily well-being' or 'healthy, looking healthy from external observation'.

7 Sino-Japanese clipping

Finally, we will take a brief look at the phenomenon of clipping in Sino-Japanese words. As in native and foreign words, clipping occurs frequently in S-J words. The original positions of the elements remaining after clipping vary greatly, as shown in (39).

(39) a. 模擬-試験 **mo**gi-**si**ken 'practice test' → 模試 *mo-si*

   b. 高等-学校 **koo**too-gak**koo** 'high school' → 高校 *koo-koo*

   c. 自宅-浪人 zi**taku-roo**nin 'home-ronin, study on one's own for another chance at university entrance exams' → 宅浪 *taku-roo*

   d. 携帯-電話 **keitai**-denwa 'mobile phone' → 携帯 *keitai*

The elements remaining after clipping can also undergo phonetic changes, as shown in (40). In (40a) [h] has changed to [p] following the moraic nasal and in (40b) vowel deletion has yielded a geminate consonant cluster.

(40) a. 民間-放送 **min**kan-**hoo**soo 'commercial broadcasting' → 民放 *min-poo*

   b. 特別-急行 **toku**betu-**kyuu**koo 'special express' → 特急 *tok-kyuu*

Unlike clipping in foreign words (cf. Chapter 5 [Irwin, this volume]), which relies on the phonetic unit of the mora, clipping in Sino-Japanese words occurs the orthographic unit of the *kanji*. That it is the *kanji* orthographical unit that is important is also suggested by pre- and post-clipping differences in the readings of the *kanji*. For example, in (41a) the second character (阪 *saka*) of 大阪 *Oosaka* 'Osaka' changes its reading to *han* (its *on*-reading) and in (41b) the first character (落 *raku*) of 落語 *rakugo* 'rakugo, comic storytelling' changes its reading to *oti* (its *kun*-reading).

(41) a. 大阪-大学 *Oosaka-daigaku* 'Osaka University' → 阪大 *han-dai*

   b. 落語-研究会 *rakugo-kenkyuukai* 'rakugo research club' → 落研 *oti-ken*

Changing the readings of *kanji* in this way can be thought of as a response to cultural and social demands for novelty and uniqueness rather than to any purely linguistic demands.

8 Conclusions and future research perspectives

This chapter has considered the word formation processes and productivity of Sino-Japanese words, categorizing them into one-, two-, three-, and four-character words. Among the reasons S-J words are so successfully productive is that there is the visual factor that the meanings of the component morphemes are easily comprehended through the *kanji*, but a more important reason is that, as long as the head is an S-J morpheme, in long compounds restrictions on lexical types are relaxed and in addition to S-J words, native and foreign morphemes may be incorporated as non-head components. Particularly noticeable are the iterative application of compounding rules to produce compounds four or more characters in length and the vigor of affixes that can attach to bases of three or more characters.

While traditional research on Sino-Japanese words in Japan has centered on discussion of the internal structure of complex words and the semantic relations among the constituents of the word based on the number of characters, in theory-oriented approaches a series of studies (Kageyama 1980, 1982, 1993) have incorporated the research methodologies of Western linguistics and examined the links between word-internal structure and syntactic structure, and other studies on lexical semantics applying Lexical Conceptual Structure and Qualia Structure have gained attention. Even more recently, studies examining the structure of Sino-Japanese words from a syntactic perspective (Ishii 2007, 2013) and studies attempting to describe the ambiguity of S-J affixes using methods from cognitive semantics (Yamashita 2011, 2013) have appeared.

An outstanding problem is how compounding and affixation in Sino-Japanese words can be differentiated. To solve these issues, it will be necessary to utilize large-scale corpora and analyze in detail the meaning and use of individual compounds and affixed words. Since research until now has concentrated on a comparatively small number of affixes, it is important to broaden the scope of analysis. Another important issue is why Sino-Japanese word formation enjoys such high productivity. This issue should be illuminated through comparison of word formation processes in the other lexical strata.

Finally, we should not forget to mention the nature of *kanji* itself. Although writing has been treated only mariginally in Western linguistics, from the practical and psychological perspectives *kanji* play an important role in providing the readers of written Japanese with a visual aid for capturing the meaning of a word at a glance. From the perspective of theoretical morphology as well, the two ways of reading

kanji have potential to shed new light on how the notions of morpheme and allo-morphy are related to differences of lexical strata. In fact, Nagano and Shimada (2014), in line with earlier studies like Morioka (2004), argue that the two different readings of a single *kanji* character should be regarded as suppletive allomorphs of one and the same underlying morpheme. The feasibility of the suppletion theory will have to be tested against a number of empirical issues involving the systematic and not so systematic correspondences between *on-yomi* (Chinese reading) and *kun-yomi* (Japanese reading).

Acknowledgments

References

Irwin, Mark. this volume. Chapter 5: The morphology of English loanwords. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Ishii, Masahiko. 2007. *Gendai nihongo no fukugōgo keiseiron* [A theory of compound word formation in contemporary Japanese]. Tokyo: Hituzi Syobo.

Ishii, Masahiko. 2013. Rinjiteki-na yoji-kango no keisei [Formation of four-character Sino-Japanese nonce words]. In Masa-aki Nomura (ed.), *Gendai nihon kango no tankyū* [Investigations into modern Sino-Japanese words], 146–166. Tokyo: Tōkyōdō Shuppan.

Kageyama, Taro. 1980. *Nichi ei hikaku goi no kōzō* [Lexical structures: English-Japanese comparison]. Tokyo: Shōhakusha.

Kageyama, Taro. 1982. Word Formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1999. *Keitairon to imi* [Morphology and semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word forma tion*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Lexical strata and word formation pro-cesses. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kobayashi, Hideki. 2004. *Gendai nihongo no kango dōmeishi no kenkyū* [A study of Sino-Japanese verbal nouns in present-day Japanese]. Tokyo: Hituzi Syobo.

Miyake, Tomohiro. 2010. Ichiji-kango *suru*-gata dōshi o megutte [On *suru*-verbs with one-character Sino-Japanese words]. In Hiroko Ōshima, Akiko Nakajima, and Raoul Blin (eds.), *Kango no gengogaku* [The linguistics of Sino-Japanese words], 107–119. Tokyo: Kurosio Publishers.

Morioka, Kenji. 2004. *Nihongo to kanji* [Japanese and *kanji*]. Tokyo: Meiji Shoin.

Nagano, Akiko and Masaharu Shimada. 2014. Morphological theory and orthography: *Kanji* as a representation of lexemes. *Journal of Linguistics* 50. 323–364.

Namiki, Takayasu and Taro Kageyama. this volume. Chapter 6: Word structure and headedness. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Nomura, Masa-aki. 1973. Hitei no settōgo *mu*-, *hu*-, *mi*-, *hi*- no yōhō [Usage of the negative prefixes *mu*-, *hu*-, *mi*-, and *hi*-]. *Kokuritsu Kokugo Kenkyūjo ronshū: Kotoba no kenkyū* [Kokuritsu Kokugo Kenkyūjo reports: Studies on language] 4. 31–50. Tokyo: Shūei Shuppan. Downloadable from http://db3.ninjal.ac.jp/publication db/

Nomura, Masa-aki. 1974. Sanji-kango no kōzō [The structure of three-character Sino-Japanese words]. *Densanki ni yoru kokugo kenkyū* [Studies on the Japanese language with computers] 6. 37–62. Tokyo: Kokuritsu Kokugo Kenkyūjo. Downloadable from http://db3.ninjal.ac.jp/publication db/

Nomura, Masa-aki. 1975. Yoji-kango no kōzō [The structure of four-character Sino-Japanese words]. *Densanki ni yoru kokugo kenkyū* [Studies on the Japanese language with computers] 7. 36–80. Tokyo: Kokuritsu Kokugo Kenkyūjo. Downloadable from http://db3.ninjal.ac.jp/publication db/

Nomura, Masa-aki. 1977. Zōgo-hō [Word-building]. *Iwanami kōza Nihongo 9: Goi to imi* [Iwanami course in Japanese 9: Meaning and vocabulary], 255–289. Tokyo: Iwanami Shoten.

Nomura, Masa-aki. 1984. Goshu to zōgoryoku [Lexical classes and productivity]. *Nihongogaku* 3(9). 40–54. Tokyo: Meiji Shoin.

Nomura, Masa-aki. 1988. Niji-kango no kōzō [The structure of two-character Sino-Japanese words]. *Nihongogaku* 7(5). 44–55. Tokyo: Meiji Shoin.

Ōtsuki, Fumihiko. 1882. *Genkai* [Wordbook]. Tokyo: Rokugōkan.

Selkirk, Elizabeth. 1982. *The syntax of words*. Cambridge, MA: MIT Press.

Shimamura, Reiko. 1985. Fukugōgo to haseigo: Kango-kei fukugōdōshi o chūshin ni [Compounds and derived words, with reference to Sino-Japanese compound verbs]. *Journal of Tsuda College* 17. 289–301.

Yamashita, Kiyo. 2011. Jion-setsubiji *-siki*, *-fuu*, *-teki* no imi: Purototaipu to sukīma [The meaning of Sino-Japanese suffixes *-siki*, *-fuu*, and *-teki*: Prototype and schema]. *Aoyama Gobun* [Aoyama Language and Literature] 41. 130–142. Tokyo: Aoyama Gakuin University.

Yamashita, Kiyo. 2013. Setsuji-sei jionkeitaiso no zōgo-kinō [The word-building function of affix-like Sino-Chinese morphemes]. In Masa-aki Nomura (ed.), *Gendai nihon kango no tankyū* [Investigations into modern Sino-Japanese words], 83–108. Tokyo: Tōkyōdō Shuppan.

Yamada, Yoshio. 1940. *Kokugo ni okeru kango no kenkyū* [A study on Sino-Japanese words in Japanese]. Tokyo: Hōbunkan.

Zhu, Jingwei. 2011. Rangaku-shiryō no sanji-kango nitsuite no kōsatsu [Three-character Sino-Japanese words in written documents of Dutch studies]. *NINJAL Project Review* 4. 1–25. Tokyo: National Institute for Japanese Language and Linguistics.

Zhu, Jingwei. 2013. Chūgokugo ni shakuyō-sareta Meiji-ki no kango [The Sino-Japanese words of the Meiji period that entered Chninese]. In Masa-aki Nomura (ed.), *Gendai nihon kango no tankyū* [Investigations into modern Sino-Japanese words], 317–346. Tokyo: Tōkyōdō Shuppan.

Kimi Akita and Natsuko Tsujimura

4 Mimetics

1 Introduction

This chapter surveys the morphological, semantic, and syntactic properties of mimetic words in some detail (see Chapter 1 [Kageyama and Saito, this volume] for their general features and Nasu (2015) for their phonological properties). The term "mimetic" here maximally covers a variety of words that not only imitate the sound associated with an entity or action ("onomatopoeia") but also express manners of action/event or internal feelings and perceptual sensations associated with an entity or action.

Mimetic lexemes across languages have been a center of debate regarding their linguistic significance. On the one hand, in European languages that have a relatively small inventory of this lexical class, it is often believed that mimetic words show largely idiosyncratic characteristics that do not contribute much to the linguistic pursuit of the nature of human language. Representative mimetic words in English, for example, include words that mimic sounds (e.g. *bow-wow, tick-tock, buzz*, *purr*) and those that do not exactly imitate sounds (e.g. *helter-skelter, flip-flop*, *twinkle*), but their number and grammatical usage are extremely limited. On the other hand, a number of languages of the world including Japanese, Korean, African languages (e.g. Hausa, Chichewa), South and Southeast Asian languages (e.g. Tamil, Semai), and South American languages (e.g. Quechua) have an enormously rich repertoire of mimetic words that are used with high frequency and with unique grammatical properties (Hinton, Nichols, and Ohala (eds.) 1994; Sohn 1994; Nuckolls 1996; Voeltz and Kilian-Hatz (eds.) 2001; Ibarretxe-Antuñano 2006).[1]

In these languages, the number of mimetic words and the extent to which they are used in communication are significant enough to warrant treating them as no different from garden-variety words like nouns and verbs. Extensive research for the past fifty years or so has revealed that mimetic words in Japanese show a number of interesting characteristics that are worth serious linguistic investigation. For example, a large inventory of mimetic words in Japanese take numerous morphophonological forms that vary in length and phonotactics, as is demonstrated by the examples in (1). Notes on representation: "N" in mimetic words stands for a coda nasal, and "Q" for the first part of a geminate consonant when it occurs stem-medially (e.g. *goQsori* [gos:ori]) or followed by a quotative marker (e.g. *pakapakaQ to* [pakapakat:o]); otherwise /Q/ refers to a glottal stop (e.g. *paaQ* [pa:ʔ]).

---

1 Mimetic, sound symbolic lexemes are more commonly called "ideophones" in African linguistics and "expressives" in South and Southeast Asian linguistics.

(1) *piN* [pin], *kiQ* [kiʔ], *paaQ* [pa:ʔ], *kitiN* [kitʃin], *garaN* [gaɾan], *garari* [gaɾaɾi], *baaN* [ba:n], *pitiQ* [pitʃiʔ], *pesyari* [peʃaɾi], *kurukuru* [kɯɾɯkɯɾɯ], *dosuNdosuN* [dosɯndosɯn], *goQsori* [gos:ori], *dotabata* [dotabata], *huNwari* [ɸɯnwaɾi], *poQtyari* [pot:ʃaɾi]

Furthermore, mimetic words can exercise different syntactic functions when put in phrases and sentences, ranging over nouns, adjectivals, adverbials, and verbs, as exemplified by *hirahira* (representing a 'fluttering' or 'flapping' state) in (2).[2]

(2) a. **Hirahira** *ga* *kininaru.* (noun)
MIM NOM be.conscious
'(He) is conscious about the flapping object.'

b. **hirahira** *no/na* *sukaato* (adjectival)
MIM GEN/COP skirt
'fluttering (flare) skirt'

c. *Sakura* *no* *hanabira* *ga* **hirahira** **to** *tiru.* (adverbial)
cherry GEN petal NOM MIM QUOT fall
'Cherry petals fall in a fluttering manner.'

d. *Hata* *ga* **hirahira-suru.** (verb)
flag NOM MIM-do
'A flag flutters.'

The archetypal syntactic function of a large number of Japanese mimetics is to serve as manner adverbs, as shown in (2c). Mimetic words of this function are often followed by the quotative particle *to* (or *te* in casual speech), some obligatorily and others optionally, as illustrated in (3), taken from Hamano (1998: 13) (see also Nasu 2002; Toratani 2006).

(3) a. Both *to* and *te* possible: *pipiN to/te, piiN to/te, pakapakaQ to/te, pataQ pataQ to/te*

b. obligatory *to*: *paQ-pa to*

c. optional *to*: *piNpiN (to), piipii (to), pakupaku (to), piQtari (to), uNzari (to)*

---

2 The mimetic word *hirahira* in (2b) is "adjectival" in the sense that it modifies the noun *sukaato* 'skirt'. Based on morphological criteria, however, *hirahira* in this example is either a noun or an adjectival noun: as a noun, *hirahira* takes the Genitive Case particle *no* following the prenominal modification pattern specific for nouns; as an adjectival noun, *hirahira* appears in its prenominal inflectional pattern with *na*. *Hirahira* is more frequently used as *hiranira no* + N, but the particular example of *hirahira na sukaato* was found in our internet search.

From a semantic perspective, Japanese mimetics have traditionally been divided into three classes according to their descriptive nature: phonomimes (*giongo/giseigo*), phenomimes (*gitaigo*), and psychomimes (*gijōgo*) (Martin 1975). Phonomimes mimic sounds of animate and inanimate beings, phenomimes describe senses that are non-auditory such as visual and tactile senses, and psychomimes are used to depict emotional, psychological states, or bodily sensations. The three classes are exemplified in (4).

(4) a. Phonomimes: *nyaa* 'meow', *dosiN* 'thud', *bakiQ* 'crack', *kusukusu* 'chuckle'

   b. Phenomimes: *siQtori* 'moist', *niyari* 'grinning', *nebaneba* 'sticky', *tyokomaka* 'bustling restlessly'

   c. Psychomimes: *biQkuri* 'surprised', *kuyokuyo* 'worrying', *tikuQ* 'prickled'

Investigations of each aspect of these and other properties have elucidated the degree of productivity and of creativity of mimetics in comparison with words in the other strata such as *Yamato kotoba* or *wago* (native Japanese words), Sino-Japanese (loanwords from Chinese), and foreign (loanwords from languages other than Chinese). What has been found and discussed in exploration of Japanese mimetics has contributed to a renewed way in which this lexical class (vocabulary stratum) is viewed vis-à-vis descriptive, typological, and theoretical issues that have been relevant to non-mimetic words. Extensive inquiry has been made in the world's languages with a large set of mimetic words at a similar level of intensity and rigor. In these exploratory efforts, it has been revealed that this lexical class plays an important role in linguistic research in at least two respects. First, what appears to be idiosyncratic in comparison with other lexical classes (i.e. strata) may turn out to form a set of properties that sums up to regularity within the mimetic lexical class, and the cluster of properties specific to the mimetic class may reveal an aspect of human language that other lexical classes are not able to disclose. Second, the manner in which mimetic lexemes interact with other classes of the lexicon enhances our understanding of how separate classes of words negotiate their differences in interaction and to what extent various theories and approaches that have been available primarily to non-mimetic words can be made applicable or adjusted to the mimetic class. There is no denying that linguistic research on mimetics has been fruitful in underscoring their intrinsic properties that are interesting and important in their own right and has made a contribution to linguistic theory.

In this chapter we will first give a brief summary of previous studies on mimetics, focusing on the past fifty years, primarily in Japanese but in other languages as well where relevant. The review of the past literature is intended to give a context in which more current work has been developed and elaborated on; some of it is

discussed in Section 3. This chapter will conclude by laying out the direction of future investigations and their challenges. As we have briefly suggested at the outset, research on mimetic lexemes has been extending its scope of examination, and a single chapter will not do justice to every aspect of this exciting topic. Instead, throughout our discussion below, we will concentrate mostly on the relation between form and meaning that mimetic words display, particularly from the perspective of productivity and creativity.

2 Form and meaning of mimetics

This section will provide a brief survey of the basic properties of mimetics in their phonology, morphology, semantics, and syntax.

2.1 Phonosemantics, phonology, and morphology

As is common with research of sound-symbolic words in the world's languages (e.g. Samarin 1971; Diffloth 1976; Kunene 1978; Zwicky and Pullum 1987; Hinton, Nichols, and Ohala (eds.) 1994; Voeltz and Kilian-Hatz (eds.) 2001), much attention has been paid to the extrinsic nature of this lexical class, focusing largely on descriptive generalizations of its idiosyncratic properties. Investigations of Japanese mimetics are no exception to this trend. In fact, more than twenty dictionaries have been published for this particular lexical class not only in Japanese (e.g. Asano (ed.) 1978; Atōda and Hoshino 1995; Ono (ed.) 2007) but also in English (Kakehi, Tamori, and Schourup 1996). The individual sets of characteristics that have been described to date are indeed interesting in their own right, but the question as to whether they have potential to contribute to generalizations pertinent to other lexical classes or to the lexicon as a whole in more general terms has not been sufficiently addressed. In particular, the morphophonological characteristics of mimetics have been typically linked up to reduplication (e.g. *huwahuwa* 'fluffy'), suffixation (e.g. *huwaQ*, *huwaN*, *huwari* 'fluffy'), and prosodic operations such as gemination (e.g. *huwaQhuwa* 'very fluffy') and vowel lengthening (e.g. *huwaaQ* 'very fluffy') (Kobayashi 1935; Izumi 1976; Kindaichi 1978; Tamamura 1984; Nishio 1988; Otsubo 1989; Kakehi and Tamori (eds.) 1993; Tamori and Schourup 1999; Yamaguchi 2002). Of relevance is that these common morphophonological characteristics show semantic repercussions. For example, reduplicative forms imply repeated or continuous events while the suffixation of *-ri* to a mimetic root may have the opposite semantic effect of quiet termination of an event; and gemination and vowel lengthening often serve as intensifiers. Similarly, it has been pointed out that mimetic pairs like *kirakira* 'glittering' vs. *giragira* 'glittering unpleasantly' and *sakusaku* 'cutting smoothly' vs. *zakuzaku* 'cutting forcefully'

suggest that voicing of initial obstruents can add "heaviness/massiveness" to the meanings of forms with voiceless counterparts (Hamano 1998; Suzuki 1965; Murakami 1980).

Apart from these descriptive studies, theoretically-oriented researchers have started to look at mimetics as forming a lexical class in which a coherent regularity can be defined as opposed to the previous assumption that the linguistic behavior of its members is arbitrary and idiosyncratic. According to this view, mimetics form a lexical class (or lexical stratum) that is on par with other lexical strata such as *Yamato* (i.e. native) words, Sino-Japanese words, and foreign words in that its members are bound by the same set of rules and constraints (McCawley 1968; Itô and Mester 1995). In the forefront of work that marks this new wave of investigations is Hamano's extensive work on the topic (Hamano 1986, 1998). She closely investigated the phoneme-meaning relations between the two consonants in C_1VC_2V-based mimetic adverbs, as is summarized in (5) and (6) below, and drew the generalization that C_1 describes the tactile nature of an object and C_2 refers to the type of movement. For example, the semantic distinction between *kotukotu* and *tokutoku* is explained as: "the combination /k-t-/ means that 'a hard surface is involved in hitting,' whereas the combination /t-k-/ means that 'a lax surface is involved in an inward/outward movement'" (Hamano 1998: 170).

(5) C_1 p   taut surface   light; small; fine

          b   taut surface   heavy; large; coarse

          t   lack of surface tension; subduedness   light; small; fine

          d   lack of surface tension; subduedness   heavy; large; coarse

          k   hard surface   light; small; fine

          g   hard surface   heavy; large; coarse

          s   non-viscous body; quietness   light; small; fine

          z   non-viscous body; quietness   heavy; large; coarse

          h   weakness; softness; unreliability;
             indeterminateness

          m   murkiness

          n   viscosity; stickiness; sliminess;
             sluggishness

          y   leisurely motion; swinging motion;
             unreliable motion

          w   human noise; emotional upheaval

(Hamano 1998: 172)

(6) C$_2$ p, b explosion; breaking; decisiveness

t hitting of a surface; coming into close contact; complete agreement

k opening; breaking up; swelling; expanding; puffing out; emission from inside; surfacing; in-out movement

s soft contact; friction

h breath

m ?

n bending; elasticity; unreliability; lack of force; weakness

y sound from many sources; haziness; childishness

w softness; faintness; haziness

r rolling; fluid movement

(Hamano 1998: 173)

Although a number of exceptions to this generalization are found, it is noteworthy that Hamano's detailed investigation brought to light a critical difference between monomoraic ((C)V) and bimoraic ((C$_1$)V$_1$C$_2$V$_2$) mimetic roots. Bimoraic roots (e.g. *poki* for *pokipoki*, *pokiN*, and *poQkiri* 'crunching') are more analytic and more linguistically constrained than monomoraic ones (e.g. *po* for *poQ* 'appearing suddenly, blushing', *poNpoN* 'tossing, tapping', and *poiQ* 'tossing') in that the latter do not seem to give rise to a clear phonosemantic generalization. That is, the meanings of CV-based mimetics are generally less transparent than those of CVCV-based ones, as the translations here suggest.

Itô and Mester (1995), following up on the original observation by McCawley (1968) about the lexical strata, discuss mimetics as a separate lexical stratum that is characterized by a set of phonological and phonotactic constraints, as shown in (7) (see Nasu 2015 for a detailed survey).

(7) a. Yamato   *P   *NT *DD
    b. Sino-Japanese *P   –   *DD
    c. Mimetic   –   *NT *DD
    d. Foreign   –   –   –

(Itô and Mester 1995: 820)

The constraint indicated by *P refers to the prohibition of a single occurrence of [p]; *NT refers to a constraint against a nasal consonant immediately followed by a voiceless stop, preventing *[nt], *[mp], and *[ŋk] from surfacing; and geminate voiced obstruents are barred by the constraint *DD. As (7) illustrates, each lexical stratum is characterized by a unique combination of these three constraints: the *Yamato* (i.e. native) stratum is most heavily constrained while the foreign stratum is free of any of these phonological restrictions, and the Sino-Japanese and mimetic

strata come in between. Mimetic words follow *NT and *DD, but interestingly the lack of *P accounts for the observation that one-sixth of mimetic words start with [p] (Hamano 1998). The survey of phonological constraints in (7) suggests that there is legitimate reason for treating mimetic words as a lexical class (or stratum) independent of each of *Yamato*, Sino-Japanese, and foreign strata.

2.2 Syntax and semantics of mimetic verbs

Along with research on the internal regularities within the lexical class of mimetics, attempts have been made to explore how this lexical class can be captured vis-à-vis various theoretical approaches that have primarily been based on non-mimetic words. Two examples are given here that illustrate theoretical application to the syntax and semantics of mimetic words. First, stressing that mimetic stems such as *burabura*, *gosogoso*, and *batabata* can potentially be interpreted in multiple ways, Tsujimura (2005a) demonstrates within the framework of Construction Grammar that even when they are combined by *suru* 'do' to form a verb, they take on a meaning specific to each morphological and structural construction in which the mimetic verb appears. The examples in (8) through (11), taken from Tsujimura (2005a: 147), show that while the mimetic stem *burabura* by itself does not have a concrete "meaning" – in a sense parallel to the "meaning" associated with non-mimetic words – to be represented by way of semantic decomposition of some sort, the verbal form, *burabura-suru*, picks out a more definitive meaning in conjunction with the global information obtained throughout the sentence in which it appears.

(8) *Doa   no   totte   ga   burabura-suru.*
    door   GEN knob NOM MIM-do
    'The door knob is loose.'

(9) *Taroo   ga   uti   de   burabura-site   iru.*
    Taro   NOM home LOC MIM-doing   be
    'Taro is being lazy at home.'

(10) *Taroo   ga   kooen   o   burabura-sita.*
    Taro   NOM park ACC MIM-did
    'Taro strolled leisurely in the park.'

(11) *Taroo   ga   asi   o   burabura-suru/burabura-s-aseru.*
    Taro   NOM leg ACC MIM-do/MIM-do-CAUS
    'Taro swings (his) legs.'

The information necessary to construe the meaning of these sentences includes the argument structure patterns coupled with the semantic and pragmatic nature of

arguments, such as the animacy of the subjects in (8) and (9) and the semantic roles of the accusative-marked arguments in (10) and (11), the morphological shape of *suru*, such as the causative morphology in (11), and the nature of adjuncts. Thus, the intransitive argument structure of (8) with an inanimate subject leads to the description of a loose doorknob; in contrast, in (11), by virtue of the presence of the object coupled with the agentive subject, the sentence is interpreted as a causative motion.

Tsujimura's Construction Grammar approach to the multiple interpretations of mimetic verbs is challenged by Kageyama (2007), who argues that alleged peculiarities of the syntactic and semantic behavior of (conventional) mimetic words can be properly assimilated to standard frameworks of lexical semantics. Kageyama first divides the meanings of mimetic verbs into seven types, as in (12), to demonstrate that the meaning of mimetic verbs is fully represented by a mechanism making use of Lexical Conceptual Structure (LCS) and that the syntactic realization of their arguments is fully predicted by general principles of linking. He then posits seven LCS templates associated with the verb *suru* as in (13). The LCS template is conflated with one of the LCS contents of mimetic bases in (14) to derive the meaning of mimetic verbs, as is schematized in (15). In this approach, the meaning of the mimetic verb *akuseku-suru*, for example, is composed through a process that combines two sets of LCS, as schematically illustrated in (16).

(12) a. Group A (Agent/Experiencer subjects)
          Type 1: [activity] (e.g. *akuseku-suru* 'work hard')
          Type 2: [impact] (e.g. *doNdoN-suru* 'bang (the wall)')
          Type 3: [manner-of-motion] (e.g. *urouro-suru* 'wander aimlessly')
          Type 4: [psychological] (e.g. *gaQkari-suru* 'be disappointed')

      b. Group B (Theme subjects)
          Type 5: [physiological] (e.g. *zukizuki-suru* 'throb (with pain)')
          Type 6: [physical perception] (e.g. *guragura-suru* 'wobble')
          Type 7: [characterizing predication] (e.g. *aQsari-suru* 'become simple and light')

(13) Lexical Conceptual Structure (LCS) templates for *suru*
      Type 1: [$_{EVENT}$ x ACT]
      Type 2: [$_{EVENT}$ x ACT]
      Type 3: [$_{EVENT}$ x CONTROL [...]]
      Type 4: [$_{EVENT}$ x EXPERIENCE [...]]
      Type 5: [$_{STATE}$ x COGNIZE [$_{EVENT}$ ...]]
      Type 6: [$_{STATE}$ x COGNIZE [$_{EVENT}$ ...]]
      Type 7: [$_{STATE}$ x COGNIZE [$_{EVENT}$ ...]]

(14) Lexical Conceptual Structure (LCS) contents for mimetic base

Type 1: Manner α of action

Type 2: $\text{ON}_{<\text{Manner } \alpha>}$ y

Type 3: x $\text{MOVE}_{<\text{Manner } \alpha>}$ [$_{\text{Route}}$]

Type 4: [$_{\text{EVENT}}$ BECOME [$_{\text{STATE}}$ x BE AT-[$\text{PSYCH.STATE}_{<\text{Manner } \alpha>}$ ABOUT-z]]]

Type 5: [$_{\text{EVENT}}$ x's BODY-PART $\text{MOVE}_{<\text{Manner } \alpha>}$]

Type 6: [$_{\text{EVENT}}$ y $\text{MOVE}_{<\text{Manner } \alpha>}$]

Type 7: [$_{\text{EVENT}}$ BECOME [y BE AT-[$\text{STATE}_{<\text{Manner } \alpha>}$]]]

(15) *suru*'s LCS template: [$_{\text{EVENT}}$ x ACT]

$\uparrow$ ---- <Manner α>: LCS content of a mimetic word

(16) a. LCS template of *suru*: [$_{\text{EVENT}}$ x ACT]

b. LCS content of *akuseku*: <Manner: BUSILY>

c. $\rightarrow$ Semantic incorporation
   akuseku-suru: [$_{\text{EVENT}}$ x $\text{ACT}_{<\text{BUSILY}>}$]

The verb *suru* is associated with the template of (16a), and incorporates the LCS content of the mimetic base of (16b); the result of the incorporation of the two sets of LCS emerges as the LCS in (16c) – i.e. the meaning representation – of the mimetic verb, *akuseku-suru*. Kageyama claims that the verb *suru* that takes part of a mimetic verb is associated with six other types of LCS templates, and that the flexibility in meaning is in part attributed to the variety of types of LCS templates that *suru* is associated with. Thus, the diversity of syntactic realization of arguments and adjuncts that is observed in (8)–(11), for instance, does not result from the elusiveness of a mimetic stem but from multiple association patterns between the LCS template of *suru* and the LCS content of a mimetic base. It is of note that the analysis that involves LCS templates and LCS contents is concerned primarily with the productivity of mimetic verbs and their corresponding meaning, although the nature of creativity may not be directly addressed.

3 Functional accounts

The growing significance of investigations of mimetic words can be seen by the extent to which they contribute to the development of linguistic approaches and theories as well as to the exploration of typological generalizations. Our discussion in this section will demonstrate that some of the productive and creative properties

of mimetics can be accommodated by theoretical frames. Here, our discussion focuses on the way in which various aspects of Construction Grammar may effectively provide theoretical tools to account for both productive and creative properties. We will further discuss crosslinguistic observations pertinent to mimetic expressions from which an interesting set of typological generalizations might be drawn.

3.1 Roles of constructions

The importance of grammatical constructions has been widely argued for in theoretical analyses of sentences, phrases, and words (e.g. Fillmore and Kay 1995; Goldberg 1995; Fried and Östman 2004). Constructions are defined as inseparable form-meaning-function complexes, which are the basic unit of analysis and representation. In Construction Grammar, linguistic expressions are not derived from their underlying forms, but they instantiate existing constructional schemas. This basic tenet of Construction Grammar has been applied to analyzing the relationship between the formal and semantic properties that mimetics show, ranging from the word level to the sentence level.

Our first example comes from a constructional perspective held at the morphological level, i.e. the word formation of mimetics, to illustrate that this approach can capture the systematic patterning between morphophonological shapes and specific interpretations associated with them (Akita 2009). Consider the examples in (17), where the mimetic bound root *boki* gives rise to radically different meanings, depending on how it is morphologically or syntactically augmented with other elements.

(17) a. *Taroo*   *ga*   *eda*   *o*   *bókiboki*   *(to)*   *otta*.
Taro   NOM   branch   ACC   MIM   QUOT   broke
'Taro broke branches with a forceful snap.'

    b. *Taroo*   *ga*   *eda*   *o*   *bokiboki*   *ni*   *otta*.
Taro   NOM   branch   ACC   MIM   COP   broke
'Taro broke a branch into pieces with force.'

    c. *Taroo*   *ga*   *eda*   *o*   *{bokiQ/bokiN/bokiri}*   *to*   *otta*.
Taro   NOM   branch   ACC   MIM   QUOT   broke
'Taro broke a branch with a crunch.'

    d. *Taroo*   *ga*   *eda*   *o*   *boQkiri*   *(to)*   *otta*.
Taro   NOM   branch   ACC   MIM   QUOT   broke
'Taro broke a branch completely.'

As the English translations imply, the mimetic words sharing the same root but appearing in different morphophonological realizations differ in aspectual properties.

The reduplicative mimetics in (17a) and (17b) have durative aspect. The "suffixal" mimetics, such as ones with -*Q*, -*N*, or -*ri* in (17c), denote a punctual event. As Hamano (1998: 106) notes, the geminate consonant -*Q* "indicates that the movement is carried out forcefully or vigorously in a single direction"; moraic nasal -*N* "indicates that the action involves elastic objects or is accompanied by a reverberation"; and -*ri* "indicates quiet ending of the movement". The mimetic *boQkiri* in (17d) illustrates what is sometimes called an "emphatic/intensified adverb" (Kuroda 1979), which takes a form consisting of a moraic consonant {C} after the first mora and the "suffix" -*ri*.

These form-meaning pairs fit the basic tenet of Construction Morphology (Booij 2010). Each morphophonological form in (17), which can be represented as a template, is iconically associated with one or more schematic, primarily aspectual, meanings. First, the durative aspect of the reduplicative forms (i.e. $C_1(´)V_1C_2V_2$-$C_1V_1C_2V_2$) is motivated by morphological iconicity of repetition.[3] The durativity is further specified by each mimetic item as repetition, e.g. *bókiboki* 'crunching repeatedly (of stick-like objects)' in (17a), and continuation, e.g. *bokiboki* 'in pieces (of a stick-like object)' in (17b). Accentuation of reduplicative mimetics is linked with their dynamicity. Accented reduplicatives ($C_1´V_1C_2V_2$-$C_1V_1C_2V_2$) are dynamic, realized as part of an adverb or verb, whereas unaccented reduplicatives ($C_1V_1C_2V_2$-$C_1V_1C_2V_2$) are static, realized as (part of) an adjective or noun (Kageyama 2007). Second, the forms that comprise the three suffixal elements (i.e. -*Q*, -*N*, and -*ri*), schematized as $C_1V_1C_2´V_2$-*Q*, $C_1V_1C_2(´)V_2$-*N*(´), and $C_1V_1C_2´V_2$-*ri*, can be abstracted into the general suffixal construction, $C_1V_1C_2(´)V_2$-Aff(´), which is linked with punctual aspect (non-reduplicated forms for short events). Each of the three subtypes, represented as subconstructions subsumed under the general suffixal construction, is paired with a set of minor semantic features cited above. Third, the semantic properties of emphatic-adverbial mimetics, schematically represented as $C_1V_1CC_2´V_2$-*ri*, are somewhat difficult to generalize, compared to those of the other two mentioned above. One important semantic feature of emphatic mimetics, however, is their ambivalence between manner and result interpretations. For example, unlike the adverbs *bokiboki* in (17a) and *bokiQ/bokiN/bokiri* in (17c), which seem to have a strong preference for the manner interpretation, the emphatic mimetic *boQkiri* in (17d) can refer to the completely broken state of a branch, but we can also interpret it to describe the forceful manner (and sound) of breaking. In fact, unlike its relatives, *boQkiri* can be used in a result description (which also suggests the precedence of a forceful breaking event), as in (18).

---

3 We adopt the standard notation ´ for an accent nucleus, which is realized as a pitch fall. The absence of the symbol in a constructional representation means unaccented (i.e. "flat") prosody. In our discussion, accentuation will not be indicated for individual mimetic items unless relevant.

(18) *Taroo* *ga* *{boQkiri/*bokiboki/*bokiQ/*bokiN/*bokiri}* *to* *oreta*
Taro NOM MIM QUOT broke

eda *o* *mituketa.*
branch ACC found
'Taro found a (completely) broken branch.'

The constructional status of emphatic mimetics receives additional support from their greater productivity compared to their "non-emphatic" counterparts (i.e. $C_1V_1C_2´V_2$-*ri*) (Akita 2009). It is also noteworthy that about half of the emphatic mimetics do not have non-emphatic counterparts (e.g. *koQteri* 'rich (of a dish)' vs. **koteri*) (Moriyama 2002). These distributional facts are captured by the templatic construction, rather than emphatic derivation, and confirm the independent status of emphatic mimetics. The form-meaning correspondence in the representative $C_1V_1C_2V_2$-based mimetic constructions is summarized in (19).

(19) a. $C_1(´)V_1C_2V_2$-$C_1V_1C_2V_2$:   −punctual (repetitive, continuative)
          $C_1´V_1C_2V_2$-$C_1V_1C_2V_2$:   +dynamic
          $C_1V_1C_2V_2$-$C_1V_1C_2V_2$:   −dynamic (i.e. state)

   b. $C_1V_1C_2(´)V_2$-Aff(´):   +punctual
       $C_1V_1C_2´V_2$-Q: quick, intense
       $C_1V_1C_2(´)V_2$-$N(´)$:   reverberant
       $C_1V_1C_2´V_2$-*ri*:   quiet

   c. $C_1V_1CC_2´V_2$-*ri*:   manner-result ambivalence

Many of the previous studies analyze -*Q*, -*N*, and -*ri*, for instance, as derivationally added to mimetic roots (Waida 1984; Kadooka 2007). The constructional view has distinct advantages over such a derivational view. The constructions account for the fact that two mimetics with the same root may have greatly different meanings. For instance, the 'exact' meaning is available for *poQkiri*, but not for *pokipoki*, *pokiQ*, and *pokiri* 'crunching'. Similarly, *batabata* can refer to the event of many people dying one after another, but *bataN* 'thudding' and *baQtari* 'thudding, running across' cannot. Furthermore, the existence of these constructional templates guarantees the remarkable creativity of Japanese mimetics. In fact, relatively new mimetics in recent years, such as *kosikosi* 'rubbing one's head against someone' ($C_1´V_1C_2V_2$-$C_1V_1C_2V_2$), *merumeru* 'texting' ($C_1´V_1C_2V_2$-$C_1V_1C_2V_2$), *mohumohu* 'downy and fluffy' ($C_1(´)V_1C_2V_2$-$C_1V_1C_2V_2$), and *poyoN* 'bouncing lightly' ($C_1V_1C_2V_2$-$N´$), instantiate some of the constructions in (19). In this connection, it has been observed that Tanka and Haiku poetry not only makes frequent use of existing mimetics, often with new meaning assignments, but actively creates innovative mimetic words. It is interesting to note that poets' creation of new mimetic words also follow the construction patterns in (19), as is illustrated by *hoyari* 'smiling gently' (Hiroshi

Ishida) ($C_1V_1C_2´V_2$-*ri*), *kokiri* 'snapping' (Kyoko Kuriki) ($C_1V_1C_2´V_2$-*ri*), *mowamowa* 'dizzy' (Kozue Uzawa) ($C_1´V_1C_2V_2$-$C_1V_1C_2V_2$), *munyoQ* 'mumbling' (Kiyohiro Tamai) ($C_1V_1C_2´V_2$-*Q*), and *sakisaki* 'crunch-crunch' (Yukitsuna Sasaki) ($C_1´V_1C_2V_2$-$C_1V_1C_2V_2$) (cf. Iizuka Shoten editorial department (ed.) 1999; Tamori 2002).

A second example we wish to discuss, this time at the morphosyntactic level, illustrates how the creative aspect of meaning that is often unpredictable from a mimetic word in isolation is dealt with in the construction approach. In addition to *burabura-suru* in (8)–(11), we discuss a wide range of argument structure patterns and their associated meanings using *huwahuwa-suru* as an example. The mimetic base, *huwahuwa*, generally describes the state of being soft, light, or fluffy. Its verbal use with the verb *suru* 'do' (*huwahuwa(-to)-suru/-si-te-iru*) appears in Kakehi, Tamori, and Schourup (1996), the largest mimetic dictionary written in English, with two definitions: (i) to be light and soft; and (ii) to be flighty, nonchalant, frivolous, etc. These definitions are given in (20) and (21), respectively, with slightly modified examples from Kakehi, Tamori, and Schourup (1996: 340, 342).

(20) *Masyumaro   wa   huwahuwa-si-te   i-te   yawarakai…*
marshmallow   TOP   MIM-do-GER   be-GER   soft
'Marshmallows are soft and squashy…'

(21) *Kimi,   sonna   huwahuwa(-to)-si-ta   kimoti   de   wa*
you   such   MIM-QUOT-do-PST   attitude   INS   TOP

gakumon   wa   yar-e-nai.
scholarly.work   TOP   do-can-NEG
'You'll never be a scholar with such a nonchalant attitude.'

The two definitions of *huwahuwa-suru* reflected in these examples are conventional ones and commonly appear in other mimetic dictionaries. To the extent that a definite set of conventionalized meanings are assigned, lexical representations of mimetic verbs may be treated on par with those of non-mimetic lexical verbs like *taberu* 'eat' and *aruku* 'walk'. In fact, this is the approach that is adopted by Kageyama (2007), discussed in the previous section. One of the semantic characteristics of mimetic verbs, however, is the ease with which their meanings and argument structure patterns can be extended beyond the range that non-mimetic lexical verbs generally allow. Because of this characteristic, non-conventional and "creative" meanings can be added to an existing mimetic verb as long as the inherent image that the mimetic base has is extended to and sustained in what is being described (cf. Tsujimura 2009).

Argument structures and meanings that are non-conventional or novel to varying degrees are amply attested especially in casual communication outlets such as informal daily dialogue and internet sites, as exemplified by (22)–(26), taken from Tsujimura (2009).

(22) *Hantosi-hodo   mae   kara   atama   ga   huwahuwa-suru-yoona   kanzi*
   half.year-about before from head   NOM MIM-do-like   feeling
   '[I have had] the feeling since a half-year ago that my head is <huwahuwa>'

(23) *X-san   no   akatyan   totemo   kawaii-desu   yo.   Huwahuwa-site.*
   X-Ms. GEN baby   very   cute-COP   SFP   MIM-do
   Ms. X's baby is very cute; he is <huwahuwa>.'

(24) *Mausupaddo   o   tukatte   mo   kaasoru   ga   huwahuwa-suru.*
   mouse.pad   ACC use   even cursor   NOM MIM-do
   'The cursor does <huwahuwa> even though I use a mouse pad.'

(25) *Desukutoppu   o   huwahuwa-suru   kerotyan   mo   tuitemasu.*
   desktop   ACC MIM-do   Kero-chan also come.with
   'It comes with Kero-chan, which does <huwahuwa> around the desktop.'

(26) *Huwahuwa   na   nyanko   ga ...   huwahuwa   na   omotya   o*
   MIM   COP cat   NOM MIM   COP toy   ACC
   huwahuwa-suru.
   MIM-do
   'A light cat does <huwahuwa> with a light toy.'

The examples from (22) to (25) take the intransitive argument structure frame: the subject of *huwahuwa-suru* in (22)–(24) is theme, and so is the subject of (25), although it could possibly be interpreted as agent. These intransitive sentences, however, are not uniform in event type and associated meaning.

   The mimetic verb in (22) describes dizziness and the condition resulting from high fever. While dizziness is more commonly depicted by *hurahura-suru* or *kura-kura-suru*, and *huwahuwa-suru* may not seem to be an accepted descriptive term by every native speaker, *huwahuwa-suru* appears more frequently than expected on the Internet. The mimetic verb in (23), although it is used intransitively as in (22), refers to the chubbiness of a baby. We find an example in Kakehi, Tamori, and Schourup (1996) in which the softness of a new-born baby's hand is described with *huwahuwa-suru*, but the use of the mimetic verb to refer to the general physical feature of a baby (presumably based on his/her round face) may be considered an extended way of linking the mimetic word to the round image. Contrastive with the stative description in (22) and (23), (24) and (25), still in the intransitive frame, denote dynamic movements: the stationary movement described in (24) may fit one of the definitions of the mimetic base *huwahuwa* in Kakehi, Tamori, and Schourup (1996) that "the manner of floating or drifting lightly on water or in the air", but (25), accompanied by an adjunct of the path expression, *desukutoppu o* 'all over the desktop', highlights a more extensive movement than in (24).

Finally, the last example in (26) takes the transitive argument structure with the agentive subject *nyanko* 'cat' in the nominative and the object *omotya* 'a toy' in the accusative, and describes a causative event of a cat moving a toy in a light and soft (gentle) manner connected to the mimetic base *huwahuwa*. Apparently, none of the specific events depicted by the mimetic verb *huwahuwa-suru* in (22)–(26) strictly matches the dictionary listing of argument structure and semantic definitions that reflect the most widely accepted and conventionalized "meanings." Two points should be stressed here: (i) the varying departure from conventional dictionary definitions is made possible by the fluid interpretation through the connection between the rudimentary image that a given mimetic base symbolizes and the depicted scenes in specific contexts, and (ii) the range of event types in sentences involving mimetic words can be narrowed down by the argument structure pattern of a given sentence coupled with the availability of appropriate adjuncts.[4] Put differently, the creative aspect of mimetic verbs results from the combination of the imagery (in the sense of Kita (1997)) that mimetic roots induce and the argument structure types in which a given mimetic verb appears. In Construction Grammar, argument structure itself constitutes a construction (Goldberg 1995). In Japanese the information concerning argument structure properties is provided in terms of the number of NPs and the case particles that accompany them. Additionally, some adjuncts, such as path expressions in (25), can further narrow the range of potential interpretations that mimetic verbs can lead to. Thus, this approach is consistent with the construction view of verb meaning (Tsujimura 2005a, 2009) in that sentences predicated by mimetic verbs receive their semantic interpretation through global information including argument structure and adjunct distribution, each of which is often accompanied by a specific case particle or postposition, as well as verbal morphology. It should be noted that the semantic interpretation of innovative mimetics like *hoyari* and *sakisaki* mentioned earlier is given in a similar manner (i.e. by global information) complemented by the semantic property that the morpho(phono)logical templates allude to.

Construction Grammar considers morpheme, word, phrase, and sentence each to potentially represent construction (Fillmore and Kay 1995; Goldberg 1995; Fried and Östman 2004). The general approach to various phenomena regarding mimetics illustrated above is consistent with the basic tenet of Construction Grammar. Thus, Construction Grammar appears to provide a useful theoretical apparatus that accommodates both the regularity and the creativity of mimetic morphemes particularly with respect to the relationship of form and meaning.

---

4 It has been demonstrated that the flexible range of argument structure and adjunct distribution is one of the important properties that distinguish mimetics from non mimetic words. See Tsujimura (2005a, 2009, 2014) for more details.

3.2 Crosslinguistic generalizations

Recent linguistic investigations of mimetics are not limited to language-internal observation; crosslinguistic and typological research has increasingly been showing fruitful results. We will discuss two examples, with particular focus on two related theoretical issues: (i) functional generalizations of mimetic morphosyntax and (ii) the place of mimetics in Talmy's (1991, 2000) typology of motion expressions.

To begin, the lexical category of mimetics has been one of the most controversial issues. Mimetics belong to one category in some languages, such as Semai (Austroasiatic) and Yir-Yoront (Pama-Nyungan), but to more than one category in other languages, such as Japanese and Sotho (Niger-Congo). Some studies approach the issue of category by asking whether mimetics constitute a unique category of their own or are subordinate to other major categories, such as verbs and adverbs (cf. Newman 1968; Samarin 1971; Bartens 2000). On the one hand, adverbs appear to be the most common and productive category for mimetics across languages (Watson 2001). On the other hand, it has been reported that many mimetics are verbal in Zulu (Niger-Congo) (Voeltz 1971), nominal in Somali (Afroasiatic) (Dhoorre and Tosco 1998), and adjectival in Cantonese (Bodomo 2006). Japanese mimetics are realized in various categories,[5] as we already saw in (2), repeated below.

(2) a. ***Hirahira***   *ga*   *ki-ni naru.* (noun)
       MIM   NOM   be.conscious
       '(He) is conscious about the flapping object.'

    b. ***hirahira***   *no/na*   *sukaato* (adjectival)
       MIM   GEN/COP   skirt
       'fluttering (flare) skirt'

    c. *Sakura*   *no*   *hanabira*   *ga*   ***hirahira***   *to*   *tiru.* (adverbial)
       cherry   GEN   petal   NOM   MIM   QUOT   fall
       'Cherry petals fall in a fluttering manner.'

    d. *Hata*   *ga*   ***hirahira***-*suru.* (verb)
       flag   NOM   MIM-do
       'A flag flutters.'

---

5 Although the accentuation property may play some role, the categorial identification given in (2) is primarily determined by the functional role that each of the mimetic word carries as well as the morphological environment in which a mimetic base appears. That is, we assume that a mimetic stem like *hirahira* does not have its categorial identification in isolation but is assigned a more definitive role according to its morphological status and/or its function in a phrase. For example, in (2c), it modifies the verb, which suggests the adverbial role; and in (2d), it is suffixed by the light verb *suru*, and this morphological combination is considered to form a verb. We leave open the issue of whether a mimetic stem or base has an inherent or independent category.

The pattern of category assignment as well as the syntactic distribution of mimetic words seems to be predictable when we take into consideration the nature of mimetic words from the point of view of iconicity. Akita (2009), for example, makes the generalization in (27) in an attempt to capture the correlation between the iconicity of mimetics and their morphosyntactic distribution. This generalization is based on the hierarchy of iconicity of words, and at the same time follows the principle of diagrammatic iconicity, according to which "less linguistic" items are kept out of language core.

(27) The iconicity-based generalization of mimetic morphosyntax
      The more iconic a mimetic is, the more likely it is realized in the periphery
      (i.e. adjunct) or outside (i.e. interjection-like) of the clause; the less iconic
      a mimetic is, the more likely it is realized in the core of the main clause
      (i.e. predicate and its argument).
                                            (adapted from Akita 2009: 268–269)

Phonomimes, which mimic physical sounds by means of linguistic sounds, are more iconic than phenomimes and psychomimes.

(28) a. Phonomimes:
          i. *Ai   ga   geragera   {*-si-te/warat-te}   ita.*
              Ai   NOM   MIM   {do-GER/laugh-GER}   was
              'Ai was {guffawing/laughing with great guffaws}.'

          ii. *Kaze   ga   pyuupyuu   {*-si-te/hui-te}   ita.*
              wind   NOM   MIM   {-do-GER/blow-GER}   was
              'A wind was {*whistling/blowing whistlingly}.'

      b. Phenomimes:
          i. *Ai   ga   tobotobo   {*-si-te/arui-te}   ita.*
              Ai   NOM   MIM   {-do-GER/walk-GER}   was
              'Ai was {*plodding/walking ploddingly}.'

          ii. *Ai   ga   burabura   {-si-te/arui-te}   ita.*
              Ai   NOM   MIM   {-do-GER/walk-GER}   was
              'Ai was {strolling/walking strollingly}.'

      c. Psychomimes:
          i. *Ai   wa   atama   ga   zukizuki   {-si-te/?itan-de}   ita.*
              Ai   TOP   head   NOM   MIM   {-do-GER/hurt-GER}   was
              'Ai felt her head {throb/hurt throbbingly}.'

          ii. *Ai   wa   ukiuki   {-si-te/?ukare-te}   ita.*
              Ai   TOP   MIM   {-do-GER/make.merry-GER}   was
              'Ai was {light-hearted/?light-heartedly in high spirits}.'

Most phonomimes cannot form a verb (i.e. are realized as an adjunct or an interjection-like element), whereas about half of the phenomimes and most psycho-mimes can (i.e. are realized as a predicate) (cf. Toratani 2015). For instance, the acceptability judgment in (28)[6] is consistent with the generalization in (27). A similar observation has been made crosslinguistically, confirming the regularity of the role that iconicity plays in accounting for the categorial status and morphosyntactic distribution of mimetics across languages (Sohn 1994; Kilian-Hatz 1999, 2001).

A related crosslinguistic generalization could be made based on token frequency. Drawing on a relatively small set of field data in Siwu (Niger-Congo), Dingemanse (2011) proposes the generalization that frequently occurring mimetics (e.g. *ɖɔbɔrɔɔ* 'soft') tend to be morphosyntactically entrenched and thus are likely to appear in categorial forms like verbs and adjectives, while infrequent mimetics are distributed as categorially independent forms like adverbs. This generalization is supported by a quantitative analysis of Japanese mimetics provided in Akita (2013). Figure 1 below shows that out of fifty randomly chosen Japanese mimetics, those with relatively high token frequency are used as a part of mimetic verbs at a higher rate. That is, there is a positive correlation between the token frequency of mimetics and their categorial status as verbs, although not all frequent mimetics can surface as verbs with *suru* (e.g. **ahaha-to-suru* 'laugh' (273 occurrences)) and some relatively infrequent mimetics can form verbs (e.g. *guNnari-suru* 'be dispirited' (only 2 occurrences)).
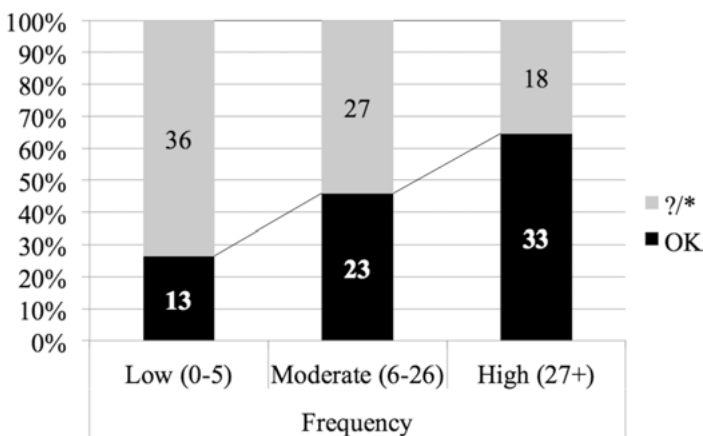

Figure 1: The frequency and verbalizability of Japanese mimetics (Akita 2013: 355)
[Note: The fifty mimetics were divided into three groups based on their relative frequency in the Balanced Corpus of Contemporary Written Japanese (NINJAL).]

---

6 The degree of acceptability of the examples in (28) seems to vary depending on the speaker, and may also depend on the contextual information that could be associated with them.

Although there are limitations and room for adjustments, the crosslinguistic data strongly suggest that iconicity and token frequency play an important role in accounting for the range of categorial status and morphosyntactic distribution of mimetic words.

Another area of crosslinguistic investigation in which mimetics prove to be of relevance is found in the lexicalization patterns or event integration. Talmy (1991, 2000) and Slobin (1996, 1997, 2000), among others, extensively discuss typological patterns in which components of motion events, including manner of motion and path of motion, are expressed in the world's languages. Speakers of "satellite-framed languages," such as Germanic, Finno-Ugric, and Slavic languages, frequently refer (and attend) to manner of motion by means of manner-of-motion verbs (e.g. *float*, *fly*, *rush*), which cannot be omitted when realized in the clause head. On the other hand, speakers of "verb-framed languages," including Romance and Semitic languages, often omit manner-of-motion information in their description of motion events, because the clause head is normally occupied by a path verb (e.g. *salir* 'exit' in Spanish) while manner is encoded in optional elements. However, as Slobin (2004, 2006) notes, the frequency of manner expressions (or "manner salience") depends not only on the two-way typology of verb-framed vs. satellite-framed languages but also on other features that are available in a given language, including richness of mimetic vocabulary and iconic gesture (see also Beavers, Levin, and Sham 2010; Toratani 2012). Slobin's observation has been further examined by some psycholinguistic methods, such as lexical comparison between originals and translations of novels (Ohara 2002; Sugiyama 2005) and speech elicitation tasks using a picture book and short videos (Akita, Matsumoto, and Ohara 2010). It has been reported in these studies that speakers of English (a satellite-framed language) are indeed more frequent users of manner expressions, particularly manner verbs such as *walk*. In contrast, speakers of Japanese (a verb-framed language) use more emphatic manner expressions by way of mimetic adverbs with expressive morphology (e.g. *piiiiiQ te tobi-mawat-te* 'flying around with a *piiiiiQ* sound') (Akita, Matsumoto, and Ohara 2010). This demonstrates that even verb-framed languages may find linguistic (and extra-linguistic) expressions that make manner of motion salient. Precisely which mechanism is to be used depends on the language, but mimetic words appear to play an important role in descriptions of motion events in languages with a rich mimetic inventory.

4 Recent developments

Recent methodological advancements in linguistics and psycholinguistics have both confirmed the validity of some previously proposed generalizations about mimetics and have led to new findings about them. Research on Japanese mimetics has been growing hand in hand with this trend in general linguistics, with corpus-based and

experimental studies at its core. Specific issues in more recent literature include the definition and collocation of mimetics and their L1 and L2 acquisition.

By virtue of the clear and often vibrant images of mimetic words that link to specific characteristics of narrowly conceived eventualities, mimetic adverbs cross-linguistically show strong collocation with constituents in the same sentence, such as verbs and subject NPs (Hirose 1981; Childs 1994; Kita 1997; Schaefer 2001; Watson 2001; Toratani 2007, 2012). For example, Kita (1997) demonstrates that the adverbial *gorogoro to*, which describes "movement of a heavy round object with continuous rotation" (p. 403), is restricted in its selection of cooccurring verbs and subjects.

(29) a. *Tetu no   tama ga   gorogoro to   {korogat-ta/*subet-ta}.*
          iron GEN ball   NOM MIM   QUOT {roll-PST/slide-PST}
          'An iron ball {rolled/*slid}.'

     b. *{Tetu no   tama/*Hitotubu no   sinzyu} ga   gorogoro to*
          {iron GEN ball/one.piece   GEN pearl}   NOM MIM   QUOT
          korogat-ta.
          roll-PST
          '{An iron ball/*A pearl} rolled on.'   (Kita 1997: 403)

The mimetic specifies the manner of motion (i.e. rolling), and therefore is not compatible with the verb *suberu* 'slide' in (29a). Moreover, the voiced initial consonant (i.e. /g/) of the mimetic specifies the rolling object as heavy and large. This explains why the cooccurrence with the NP *hitotubu no sinzyu* 'a pearl' is unacceptable while an iron ball is a suitable match for the mimetic word in (29b).

The high degree of collocation with verbs is demonstrated more extensively by thorough examinations of major corpora of novels, newspapers, and conversations (Tamaoka, Kiyama, and Miyaoka 2011; Akita 2012). It has been argued that mimetics generally enrich the manner information provided by general manner verbs (see Section 3.2). The high semantic specificity at issue also accounts for many other phenomena that have been pointed out for mimetics in the literature, such as the absence of hyponymy between mimetics (Watson 2001; Bodomo 2006; Kita 2008), the incompatibility of two mimetics in one clause (Kita 1997), the omissibility of predicates modified by mimetics (Tamori 1988, 2002), and the unavailability of short paraphrases for mimetics (Fortune 1962; Diffloth 1972). Moreover, the same semantic property suggests the uneven organization of the mimetic lexicon, which consists of several highly productive semantic clusters, such as motion, laughing, crying, tactile sensation, pain, and emotion. In contrast, Japanese appears to have no single mimetic word specifically designated for smell, taste, or color (Izumi 1976). Furthermore, the cooccurrence of adverbial mimetics and iconic gestures, which has been partially observed across languages (Childs 1994; Kita 1997), has been demonstrated by a quantitative investigation of a multimedia corpus (Son 2010). Audiovisual

records of natural conversations are what we have sought for a fuller understanding of this lexical class for highly specific eventualities.

Although the majority of research on mimetics has long been centered on adult language, children's acquisition of this lexical class is a remarkably valuable area of investigation. Earlier studies already show the richness of mimetics in child Japanese (Okubo 1967; Fernald and Morikawa 1993); the questions concerning the range of their mimetic vocabulary and the nature of their acquisition are important in and of themselves but are yet to be fully investigated. It is, furthermore, interesting to pursue the question as to whether the acquisition of mimetics could help children learn non-mimetic lexical items. Noji's (1973–1977) longitudinal speech records of a monolingual male child have been a useful data source to investigate some of these issues in some recent studies. It has been observed that Noji's son, Sumihare, started to use mimetic verbs of the form [mimetic-*suru*] but then gradually shifted to more frequent use of mimetic words as adverbs modifying verbs. For example, Sumihare used *paaN-sita* 'broke [it]' as a verb early on ([1;9]) and similar verbal patterns, many of which are unconventional in adult Japanese; but this pattern then developed to a more adult-like form [mimetic adverb + non-mimetic verb], such as *putuN to kirete* 'cut off and' later on ([2;4]) (Tsujimura 2005b; see also Murasugi and Nakatani 2013). In another quantitative investigation, it has been revealed that children do not learn all types of mimetic words randomly; instead, mimetics with higher iconicity (e.g. phonomimes) are acquired earlier than those with lower iconicity (e.g. psychomimes) (Akita 2009). This observation may have some bearing on Imai et al.'s (2008) experimental findings concerning sound symbolism (i.e. phonetic/phonological iconicity) of mimetics. They report that in a video-based forced-choice task, 3-year-olds successfully generalized novel mimetic verbs (e.g. *nosunosu-suru* 'walk heavily' (intended)) but failed to generalize novel verbs without reduplication or clear sound symbolism (e.g. *nekeru*). The results suggest that sound symbolism may help children to attend to the dynamic aspect of an event, which they have to conceptualize separately from the event participants in order to fully understand the semantic nature of verbs. Although the developmental relevance of mimetics (and iconic expressions in general) has been recognized, empirical investigations into it are still at their initial stage.

In contrast with children's high degree of access to mimetics, L2 learners of Japanese find it particularly difficult not only to understand the precise meaning of mimetic words but to put them into actual use even when they become familiar with their general meanings. This question is addressed in recent experimental studies. Iwasaki, Vinson, and Vigliocco (2007), for instance, conducted a questionnaire-based study in which English speakers were asked to rate Japanese mimetics for laughing and walking on a set of semantic-differential scales. They found that English speakers were more successful in guessing the meaning of mimetics for laughing (i.e. phonomimes) than those of mimetics for walking (i.e. phenomimes). On the other hand, English speakers failed to share the evaluative aspect of the mimetics (e.g. graceful-

ness of laughter). These results suggest that high iconicity helps not only L1 but also L2 learners of Japanese in their acquisition of mimetics, and that sound symbolism for subjective evaluation is especially arbitrary and language-specific. (cf. Oda (2000) for another experimental study of the L2 acquisition of mimetics; Osaka et al. (2003, 2004) for a series of neuropsychological studies of mimetics.) It is worth pointing out that the generalization regarding iconicity in (27) above bears intriguing implications, ranging from crosslinguistic morphosyntactic observations to language acquisition.

Finally, the longstanding issue of how to define mimetics has also been examined experimentally. Many studies discuss the absence of necessary or sufficient conditions of the membership in the mimetic category of Japanese and other languages (Newman 1968; Hamano 1998; Tamori and Schourup 1999; Dingemanse 2011). A relatively clear definition is possible for Japanese mimetics, which are primarily given by their productive morphophonological constructions discussed in Section 3.1. The role of the word-level constructions has further been experimentally confirmed.[7] For Japanese mimetics, Akita (2009) adopted a rating task that used audio-recorded mimetic words, both existing and novel, that do and do not illustrate one of the morphophonological constructions, including those in (19). This study shows that native Japanese speakers judged words with a mimetic construction (e.g. *pusipusi*) as much more mimetic than those that do not conform to the construction patterns (e.g. *metoa*). The same study additionally demonstrates that the sound-symbolic effects (voicing-size/intensity associations) of construction-satisfying words are clearer and more systematic than those that are formed outside of constructions.

5 Conclusion and future research perspectives

We will conclude this chapter by considering a few issues that future research needs to address. First, we have observed with (22–26) that mimetic verbs can potentially extend beyond their conventionalized meaning and argument structure properties, but it will be interesting to find out whether such an extension is subject to linguistically well-defined conditions. For instance, can any mimetic verb potentially receive an indefinite number of distinguishable interpretations or can it be associated with any argument structure type available to non-mimetic verbs (Toratani 2013)? And, if there is a preferred interpretation or an argument structure pattern, is it just a matter of the speaker's individual preference, or is there any underlying linguistic generalization to be captured?

---

7 Dingemanse (2011: 25) arrives at the following general definition of mimetics (or ideophones in his terminology) that is meant to apply crosslinguistically: "Ideophones are marked words that depict sensory imagery".

Second, as we discussed in Section 3.2, we need to consider Japanese mimetics within a broader context of crosslinguistic/typological comparison. Recent field linguists have uncovered the pragmatic and anthropological features of sound-symbolic words in Niger-Congo and Quechua languages. For example, Nuckolls (1996) points out the reflection of animism, among others, in Quechua ideophones, which is viewed as an instance of the cultural significance of sound-symbolic words. Moreover, Dingemanse (2011) pursues the pragmatic nature of ideophonic words based on conversation analysis in Siwu. He discusses the mode of signification of ideophones as "depiction," which is most typically illustrated by pictures and is contrasted with "description," arbitrary coding, typical of prosaic words. He also observes the "unexpected" occurrence of ideophones in some culturally burdened linguistic contexts, such as greetings and funeral dirges. Furthermore, a typological study will certainly shed light on the still unsolved question of the factors that contribute to the richness of mimetics in a language. Nuckolls (2004) observes that sound-symbolic words typically abound in relatively non-urban areas of the world, such as Africa, Southeast Asia, and the Amazon. Japanese stands as an interesting exception to this remark. Nuckolls speculatively ascribes this to the animistic culture of Japan. It is hoped that placing Japanese mimetics in typological contexts will help answer some of these essential questions about sound-symbolic words in general.

Acknowledgments

References

Akita, Kimi. 2009. *A grammar of sound symbolic words in Japanese: Theoretical approaches to iconic and lexical properties of mimetics*. Kobe: Kobe University dissertation.

Akita, Kimi. 2012. Toward a frame-semantic definition of sound-symbolic words: A collocational analysis of Japanese mimetics. *Cognitive Linguistics* 23(1). 67–90.

Akita, Kimi. 2013. Onomatope/onshōchō no kenkyūshi [A research history of mimetics and sound symbolism]. In Kazuko Shinohara and Ryoko Uno (eds.), *Onomatope kenkyū no shatei: Chika zuku oto to imi* [Sound symbolism and mimetics: Rethinking the relationship between sound and meaning in language], 333–364. Tokyo: Hituzi Syobo.

Akita, Kimi, Yo Matsumoto and Kyoko Hirose Ohara. 2010. Idō-hyōgen no ruikeiron ni okeru chokuji-teki-keiro-hyōgen to yōtai-goi-repātorī [Deictic path expressions and manner lexicon in the typology of motion expressions]. In Taro Kageyama (ed.), *Lexicon forum* 5, 1–25. Tokyo: Hituzi Syobo.

Asano, Tsuruko (ed.). 1978. *Giongo/gitaigo jiten* [A dictionary of mimetics]. Tokyo: Kadokawa Shoten.

Atōda, Toshiko and Kazuko Hoshino. 1995. *Giongo/gitaigo tsukaikata jiten* [A dictionary of the usage of mimetics]. Tokyo: Sōtakusha.

Bartens, Angela. 2000. *Ideophones and sound symbolism in Atlantic creoles*. Helsinki: Finnish Academy of Sciences and Letters.

Beavers, John, Beth Levin and Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46(2). 331–377.

Bodomo, Adams B. 2006. The structure of ideophones in African and Asian languages: The case of Dagaare and Cantonese. In John Mugane, John P. Hutchison and Dee A. Worman (eds.), *Selected Proceedings of the 35th Annual Conference on African Linguistics: African Languages and Linguistics in Broad Perspectives*, 203–213. Somerville, MA: Cascadilla Proceedings Project.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Childs, G. Tucker. 1994. African ideophones. In Leanne Hinton, Johanna Nichols and John J. Ohala (eds.), *Sound symbolism*, 178–204. Cambridge: Cambridge University Press.

Dhoorre, Cabdulqaadir Salaad and Mauro Tosco. 1998. 111 Somali ideophones. *Journal of African Cultural Studies* 11(2). 125–156.

Diffloth, Gérard. 1972. Notes on expressive meaning. In Paul M. Peranteau, Judith N. Levi and Gloria C. Phares (eds.), *Papers from the Eighth Regional Meeting Chicago Linguistic Society*, 440–447.

Diffloth, Gérard. 1976. Expressives in Semai. In Philip N. Lenner, Laurence C. Thompson and Stanley Starosta (eds.), *Austroasiatic studies*, Part I, 249–264. Honolulu: University of Hawai'i Press.

Dingemanse, Mark. 2011. *The meaning and use of ideophones in Siwu*. Nijmegen: Max Planck Institute for Psycholinguistics/Radboud University dissertation.

Fernald, Anne and Hiromi Morikawa. 1993. Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development* 64. 637–656.

Fillmore, Charles J. and Paul Kay. 1995. *Construction grammar coursebook*. Ms. University of California, Berkeley.

Fortune, George. 1962. *Ideophones in Shona: An inaugural lecture given in the University College of Rhodesia and Nyasaland on 28 April 1961*. London: Oxford University Press.

Fried, Mirjam and Jan-Ola Östman. 2004. Construction Grammar: A thumbnail sketch. In Mirjam Fried and Jan-Ola Östman (eds.), *Construction Grammar in a cross language perspective*, 11–86. Amsterdam: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

Hamano, Shoko Saito. 1986. *The sound symbolic system of Japanese*. Gainesville FD: University of Florida dissertation.

Hamano, Shoko. 1998. *The sound symbolic system of Japanese*. Standford: CSLI Publications and Tokyo: Kurosio Publishers.

Hinton, Leanne, Johanna Nichols and John J. Ohala (eds.). 1994. *Sound symbolism*. Cambridge: Cambridge University Press.

Hirose, Masayoshi. 1981. *Japanese and English contrastive lexicology: The role of Japanese "mimetic adverbs"*. Berkeley CA: University of California dissertation.

Ibarretxe-Antuñano, Iraide. 2006. *Sound symbolism and motion in Basque*. München: LINCOM Europa.

Iizuka Shoten editorial department (ed.). 1999. *Tanka no gihō: Onomatope* [Techniques of *tanka*: onomatopoeia]. Tokyo: Iizuka Shoten.

Imai, Mutsumi, Sotaro Kita, Miho Nagumo and Hiroyuki Okada. 2008. Sound symbolism facilitates early verb learning. *Cognition* 109(1). 54–65.

Itô, Junko and Armin R. Mester. 1995. Japanese phonology. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Oxford: Blackwell.

Iwasaki, Noriko, David P. Vinson and Gabriella Vigliocco. 2007. What do English speakers know about *gera gera* and *yota yota*?: A cross-linguistic investigation of mimetic words for laughing and walking. *Japanese Language Education around the Globe* 17. 53–78.

Izumi, Kunihisa. 1976. *Giseigo/gitaigo no tokushitsu* [Some characteristics of mimetics]. In Takao Suzuki (ed.), *Nihongo no goi to hyōgen* [Japanese vocabulary and expressions], 105–151. Tokyo: Taishukan Shoten.

Kadooka, Ken-ichi. 2007. *Nihongo onomatope goi ni okeru keitai teki/on'in teki tokusei ni tsuite* [On morphological and phonological systematicity of Japanese onomatopoeic words]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2007. Explorations in the conceptual semantics of mimetic verbs. In Bjarke Frellesvig, Masayoshi Shibatani and John Smith (eds.), *Current issues in the history and structure of Japanese*, 27–82. Tokyo: Kurosio Publishers.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Lexical strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kakehi, Hisao and Ikuhiro Tamori (eds.). 1993. *Onomatopia: Gion /gitaigo no rakuen* [Onomatopia: A utopia of mimetics]. Tokyo: Keisō Shobō.

Kakehi, Hisao, Ikuhiro Tamori and Lawrence Schourup. 1996. *Dictionary of iconic expressions in Japanese*. Berlin: Mouton de Gruyter.

Kilian-Hatz, Christa. 1999. *Ideophone: Eine typologische Untersuchung unter besonderer Berück sichtigung afrikanischer Sprachen* [A typological study with special reference to African languages]. Habilitationsschrift: Universität zu Köln.

Kilian-Hatz, Christa. 2001. Universality and diversity: Ideophones from Baka and Kxoe. In Friedrich K. Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*, 155–164. Amsterdam: John Benjamins.

Kindaichi, Haruhiko. 1978. Giongo/gitaigo gaisetsu [Overview of mimetics]. In Tsuruko Asano (ed.), *Giongo/gitaigo jiten* [A dictionary of mimetics], 3–25. Tokyo: Kadokawa Shoten.

Kita, Sotaro. 1997. Two-dimensional semantic analysis of Japanese mimetics. *Linguistics* 35(2). 379–415.

Kita, Sotaro. 2008. World-view of protolanguage speakers as inferred from semantics of sound symbolic words: A case of Japanese mimetics. In Nobuo Masataka (ed.), *The origins of language: Unraveling evolutionary forces*, 25–38. Tokyo: Springer.

Kobayashi, Hideo. 1935. *Kokugo shōchōgo no kenkyū* [A study of sound-symbolic words in Japanese]. *Bungaku* 1(8). 1313–1359.

Kunene, Daniel P. 1978. *The ideophone in Southern Sotho*. Berlin: Dietrich Reimer.

Kuroda, S.-Y. 1979. *Generative grammatical studies in the Japanese language*. New York: Garland.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Moriyama, Keiko. 2002. "CVQCVri"-gata no onomatope [Onomatopoeia in the CVQCVri shape]. *Journal of International Student Center* 10. 53–71. Nagasaki University.

Murakami, Yoshihiro. 1980. *Onshōchō kasetsu no kentō: Onso, SD hō, meishi oyobi dōshi no rensōgo ni yoru seibun no chūshutsu to, sorera no kurasutāka ni yoru giongo/gitaigo no bunseki* [The inquiry of phonetic symbolism hypothesis: Analyses of onomatopoeias by the clustering of the introduced components on phonemes, SD technique, associated nouns, and associated verbs]. *The Japanese Journal of Educational Psychology* 28(3). 183–191.

Murasugi, Keiko and Tomomi Nakatani. 2013. Three types of "root infinitives": Theoretical implications from child Japanese. In Bjarke Frellesvig and Peter Sells (eds.), *Japanese/Korean Linguistics* 20. 263–279. Stanford: CSLI Publications.

Nasu, Akio. 2002. *Nihongo onomatope no gokeisei to inritsu kōzō* [Word formation and prosodic structure of Japanese onomatopoeia]. Tsukuba: University of Tsukuba dissertation.

Nasu, Akio. 2015. The phonological lexicon and mimetic phonology. In Haruo Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 253–288. Berlin: Mouton de Gruyter.

Newman, Paul. 1968. Ideophones from a syntactic point of view. *Journal of West African Languages* 5. 107–117.

Nishio, Toraya. 1988. *Gendai goi no kenkyū* [A study of the vocabulary of present-day Japanese]. Tokyo: Meiji Shoin.

Noji, Junya. 1973–1977. *Yōjiki no gengo seikatsu no jittai*, I–IV [The actual conditions of infant language life, I–IV]. Hiroshima: Bunka Hyōron Shuppan.

Nuckolls, Janis B. 1996. *Sounds like life: Sound symbolic grammar, performance, and cognition in Pastaza Quechua*. New York: Oxford University Press.

Nuckolls, Janis B. 2004. To be or not to be ideophonically improverished. *Texas Linguistic Forum* 47. 131–142.

Oda, Hiromi. 2000. *An embodied semantic mechanism for mimetic words in Japanese*. Bloomington, IN: Indiana University dissertation.

Ohara, Kyoko H. 2002. Linguistic encodings of motion events in Japanese and English: A preliminary look. *The Hiyoshi Review of English Studies* 41. 122–153. Keio University.

Ōkubo, Ai. 1967. *Yōji gengo no hattatsu* [Development of infant language]. Tokyo: Tokyōdō.

Ono, Masahiro (ed.). 2007. *Giongo/gitaigo 4500: Nihongo onomatope jiten* [Mimetics 4500: A dictionary of Japanese onomatopoeia]. Tokyo: Shōgakukan.

Osaka, Naoyuki, Mariko Osaka, Hirohito Kondo, Masanao Morishita, Hidenao Fukuyama and Hiroshi Shibasaki. 2003. An emotion-based facial expression word activates laughter module in the human brain: A functional magnetic resonance imaging study. *Neuroscience Letters* 340. 127–130.

Osaka, Naoyuki, Mriko Osaka, Masanao Morishita, Hirohito Kondo and Hidenao Fukuyama. 2004. A word expressing affective pain activates the anterior cingulated cortex in the human brain: An fMRI study. *Behavioural Brain Research* 153. 123–127.

Ōtsubo, Heiji. 1989. *Giseigo no kenkyū* [Studies on mimetics]. Tokyo: Kazama Shobō.

Samarin, William J. 1971. Survey of Bantu ideophones. *African Language Studies* 12. 130–168.

Schaefer, Ronald P. 2001. Ideophonic adverbs and manner gaps in Emai. In Friedrich K. Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*, 339–354. Amsterdam: John Benjamins.

Slobin, Dan I. 1996. Two ways to travel: Verbs of motion in English and Spanish. In Masayoshi Shibatani and Sandra A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 195–219. Oxford: Oxford University Press.

Slobin, Dan I. 1997. Mind, code, and text. In Joan Bybee, John Haiman and Sandra A. Thompson (eds.), *Essays on language function and language type*, 437–467. Amsterdam: John Benjamins.

Slobin, Dan I. 2000. Verbalized events: A dynamic approach to linguistic relativity and determinism. In Susanne Niemeier and René Dirven (eds.), *Evidence for linguistic relativity*, 107–138. Amsterdam: John Benjamins.

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist and Ludo Th. Verhoeven (eds.), *Relating events in narrative: Typological and contextual perspectives*, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, Dan I. 2006. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In Maya Hickmann and Stéphane Robert (eds.), *Space in languages: Linguistic systems and cognitive categories*, 59–81. Amsterdam: John Benjamins.

Sohn, Ho-Min. 1994. *The Korean language.* Cambridge: Cambridge University Press.

Son, Youngsuk. 2010. *Giongo/gitaigo to miburi: Terebi hōsō no maruchi media kōpasu ni yoru keiryō teki bunseki* [Relation of Japanese onomatopoeias and gestures: A multimedia corpus-based quantitative study of Japanese television broadcasts]. *Mathematical Linguistics* 27(4). 131–153.

Sugiyama, Yukiko. 2005. Not all verb-framed languages are created equal: The case of Japanese. In Rebecca T. Cover and Yuni Kim (eds.), *Proceedings of the Thirty First Annual Meeting of the Berkeley Linguistics Society, General Session and Parasession on Prosodic Variation and Change*, 299–310.

Suzuki, Masako. 1965. *Mukashi no giseigo/gitaigo* [Mimetics in the old days]. *Gengo Seikatsu* 171. 60–65.

Talmy, Leonard. 1991. Path to realization: A typology of event conflation. In Laurel A. Sutton, Christopher Johnson and Ruth Shields (eds.), *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 480–519.

Talmy, Leonard. 2000. *Toward a cognitive semantics, volume II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Tamamura, Fumio. 1984. *Onshōchōgo no gokei (sono 1)* [Forms of sound-symbolic words: Part 1]. *Doshisha kokubungaku* 24. 74–79. Doshisha University.

Tamaoka, Katsuo, Sachiko Kiyama and Yayoi Miyaoka. 2011. *Shinbun to shōsetsu no kōpasu ni okeru onomatope to dōshi no kyōki patan* [Collocation patterns of sound-symbolic words and verbs in corpora of a newspaper and novels]. *Gengo Kenkyu* 139. 57–84.

Tamori, Ikuhiro. 1988. Japanese onomatopes and verbless expressions. *Jimbun Ronshu: Journal of Cultural Science* 24(2). 105–129. Kobe University of Commerce.

Tamori, Ikuhiro. 2002. *Onomatope: Gion/gitaigo o tanoshimu* [Onomatopoeia: Enjoying mimetics]. Tokyo: Iwanami Shoten.

Tamori, Ikuhiro and Lawrence Schourup. 1999. *Onomatope: Keitai to imi* [Onomatopoeia: Form and meaning]. Tokyo: Kurosio Publishers.

Toratani, Kiyoko. 2006. On the optionality of *to*-marking on reduplicated mimetics in Japanese. In Timothy J. Vance and Kimberly Jones (eds.), *Japanese/Korean Linguistics* 14, 415–422. Stanford: CSLI Publications.

Toratani, Kiyoko. 2007. An RRG analysis of manner adverbial mimetics. *Language and Linguistics* 8(1). 311–342.

Toratani, Kiyoko. 2012. The role of sound-symbolic forms in motion event descriptions: The case of Japanese. *Review of Cognitive Linguistics* 10(1). 90–132.

Toratani, Kiyoko. 2013. Constructions in RRG: A case study of mimetic verbs in Japanese. In Brian Nolan and Elke Diedrichsen (eds.), *Linking constructions into functional linguistics: The role of constructions in grammar*, 41–66. Amsterdam: John Benjamins.

Toratani, Kiyoko. 2015. Iconicity in the syntax and lexical semantics of mimetics in Japanese. In Masako K. Hiraga, Kimi Akita, William J. Herlofsky and Kazuko Shinohara (eds.), *Iconicity: East meets west*, 125–141. Amsterdam: John Benjamins.

Tsujimura, Natsuko. 2005a. A constructional approach to mimetic verbs. In Mirjam Fried and Hans C. Boas (eds.), *Grammatical constructions: Back to the roots*, 137–154. Amsterdam: John Benjamins.

Tsujimura, Natsuko. 2005b. Mimetic verbs and innovative verbs in the acquisition of Japanese. In Rebecca T. Cover and Yuni Kim (eds.), *Proceedings of the Thirty First Annual Meeting of the Berkeley Linguistics Society, General Session and Parasession on Prosodic Variation and Change*, 371–382.

Tsujimura, Natsuko. 2009. *Onomatope dōshi no imi/kō kōzō no ichi kōsatsu* [Examination of the meaning and argument structure of mimetic verbs]. *KLS* 29: *Proceedings of the Thirty Third Annual Meeting of the Kansai Linguistic Society*, 334–343.

Tsujimura, Natsuko. 2014. Mimetic verbs and meaning. In Franz Rainer, Francesco Gardani, Hans Christian Luschützky and Wolfgant U. Dressler (eds.), *Morphology and meaning*, 303–314. Amsterdam: John Benjamins.

Voeltz, Erhard Friedrich K. 1971. Toward the syntax of the ideophone in Zulu. In Chin-Wu Kim and Herbert Stahlke (eds.), *Papers in African linguistics*, 141–152. Carbondale: Linguistic Research.

Voeltz, Friedrich K. Erhard and Christa Kilian-Hatz (eds.). 2001. *Ideophones*. Amsterdam: John Benjamins.

Waida, Toshiko. 1984. English and Japanese onomatopoeic structures. *Studies in English* 36. 55–79. Osaka Women's University.

Watson, Richard L. 2001. A comparison of some Southeast Asian ideophones with some African ideophones. In Friedrich K. Erhard Voeltz and Christa Kilian-Hatz (eds.), *Ideophones*, 385–406. Amsterdam: John Benjamins.

Yamaguchi, Nakami. 2002. *Inu wa* biyo *to naite ita: Nihongo wa giongo/gitaigo ga omoshiroi* [Dogs use to cry *biyo*: It is mimetics that are interesting in Japanese]. Tokyo: Kōbunsha.

Zwicky, Arnold M. and Geoffrey K. Pullum. 1987. Plain morphology and expressive morphology. In Jon Aske, Natasha Beery, Laura Michaelis and Hana Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, General Session and Parasession on Grammar and Cognition*, 330–340.

Corpus

BCCWJ (Balance Corpus of Contemporary Written Japanese). 2011. Tokyo: National Institute for Japanese Language and Linguistics. http://www.kotonoha.gr.jp/shonagon/

Mark Irwin
5 The morphology of English loanwords

1 Introduction

The study of Western loanwords in Japanese has a long history, with scholarly articles dating to as far back as Ōtsuki (1884) and loanword dictionaries to Ueda et al. (1915). The earliest research focusing on English loans into Japanese can be traced back to Ichikawa (1929, 1930). Although Ichikawa published in English, for many decades thereafter, scholarly output was predominantly in Japanese and confined to university or research institute bulletins. International impact was low. It was not until the last quarter of the 20th century that interest took off at a global level, with scholarly articles appearing in English in journals, handbooks and conference proceedings: Quackenbush (1977), Koo and Honma (1989), Hirata (1990), Hirozane (1992), Takagi and Mann (1994), Katayama (1995). It is at this juncture also that theoretical analyses of a morphological and morphophonological nature began to appear: J. Itô (1990), Itô and Mester (1992), Kubozono (1999b, 2001b, 2002).

Although this chapter will deal specifically with English loans into Japanese, the morphological processes to be described below are applicable to loanwords (see my definition in Section 1.1) from any language. English borrowings have become so dominant (Section 1.2), however, that there is a tendency to view any loanword as having an English source. This not only means that many Japanese mistakenly think the English for 'part-time job' and 'bread' are *arbite* and *pan*,[1] but also that the morphological processes described in this chapter can only have an English donor input. This latter assumption, indeed, has given rise to the Japanese term *waseieigo* 和製英語 'English made in Japan' – applicable just as equally to loans borrowed from French, Russian, Korean or whatever – to describe a range of very different phenomena: essentially all the phenomena described in this chapter, as well as mora clipping (Section 3.1), semantic change (Irwin 2011a: 153–155), and any form of irregular phonological adaptation (Section 1.3), especially auditory and spelling loans (Irwin 2011a: 76–80). The term *waseieigo* goes back to at least Umegaki (1944: 304). Although various categorizations – some more thorough than others – have been proposed (Ishiwata 1976; Tanabe 1990; Hirai 2003; Shibasaki, Tamaoka and Takatori 2007; Suzuki 2008), the term has suffered from chronic overuse and a deleterious lack of definition in much academic literature.[2] Needless to say, I will forgo it in the discussion below.

---

1 Japanese *arubaito* from German *Arbeit* 'work' (with subsequent semantic narrowing); Japanese *paN* from Portuguese *pão* 'bread'.
2 The non academic literature is even more guilty in this respect. In particular, mass market publications aimed at the hugely lucrative English education market bandy the term around with total impunity.

Before considering the morphology of English loanwords in Sections 2 and 3, some preliminaries are necessary. I consider the definition of a loanword in Section 1.1, statistical and attestatory detail in 1.2, issues of phonological adaptation and phonemicization in 1.3, and loanword compounding in general in 1.4.

1.1 Definitions

Modern Japanese vocabulary is conventionally divided into four major strata: native (*wago* 和語, etc.), mimetic (*giongo* 擬音語, *gitaigo* 擬態語, *giseigo* 擬声語), Sino-Japanese (*kango* 漢語, etc.) and loan (*gairaigo* 外来語, *shakuyōgo* 借用語, *yōgo* 洋語).[3] A number of definitions have been proposed for the loan stratum over the years (Itô and Mester 1999; NKD 2000–2002; Sugimoto 2007: 408, *inter alia*). However, a robust definition must eschew extraneous factors such as orthography and instead focus on the following three factors:

CHRONOLOGY: Sino-Japanese is also, technically, a stratum comprised mostly of borrowings, taken exclusively from Chinese. Crucially, however, it is felt by Japanese speakers to be different from the loan stratum and exhibits very different phonotactics. Sino-Japanese borrowing was largely complete by the 13th century.[4] The Japanese language first came into contact with Western languages in 1543, when Portuguese sailors landed on Tanegashima in the Ōsumi islands (now Kagoshima prefecture).

PHONOLOGICAL ADAPTATION: A prerequisite for loanword status is that all donor phones be assimilated into Japanese phonology (Section 1.3). As the Japanese phonemic inventory is comparatively impoverished and its phonotactics comparatively constrained, phonological adaptation can be severe. Loans whose phones remain unassimilated and pronounced 'in a foreign manner' must be considered *gaikokugo* 外国語 'foreign words'.[5]

INTELLIGIBILITY: Any loanword whose meaning is unintelligible to the general Japanese speaker community may, arguably, be considered not to have been fully incorporated into the lexicon. The intelligibility factor is, however, highly problematic. Although a soccer enthusiast may be *au fait* with the technical term *ofusaido+toraQpu*

---

3 The Japanese terms for the loan stratum have subtle differences in meaning: see Irwin (2011a: 7). Non stratal models have been proposed by Itô and Mester (1995, 1999) and Fukuzawa and Kitahara (2005).

4 For more detail on the history of Sino Japanese borrowing see *inter alia* Tōdō (1969); Frellesvig (2010: 258 294); Yuzawa (2010).

5 For a detailed discussion of phonological adaptation see Irwin (2011a: 71 136). In all the examples to follow, borrowed English words (English donor etyma) are given a dotted underline, → indicates 'adapted to Japanese', and ← 'adapted from English'. Thus, 'screen → *sukuriiɴ*' means 'English *screen* is adapted to Japanese *sukuriiɴ*'. Element boundaries in compounds are indicated with a plus sign.

(← offside trap), what proportion of the speaker community must fail to comprehend a loanword for it to be generally unintelligible? This argument applies not only to technical terms, but also to loanwords used by particular age cohorts, ethnic groups, dialect speakers, or socio-economic groups. It must also be remembered that it is not uncommon for loans to have been intelligible in the past but be unintelligible in the present, while many loans intelligible at the time of writing will fall out of use and be rendered unintelligible to future Japanese speakers.

I thus follow Irwin (2011a: 10) and define a Japanese loanword as:

(1)   a foreign word which has undergone adaptation to Japanese phonology,
        has been borrowed into Japanese after the mid-16th century and whose
        meaning is, or has been, intelligible to the general speech community.

1.2 Statistics and attestations

Loanwords make up 2%–12% of the Japanese lexicon by token, and 5%–35% by type, with great variation apparent across media.[6] Borrowings from English currently comprise the vast bulk, although this has not always been the case. In the late 19th and very early 20th century English donor words accounted for just over half of all loans, though this proportion had grown to around 75% by the 1950s. For the last half century, since the 1960s, it has hovered virtually unchanged at 80%–85% by type, or 85%–90% by token (Irwin 2011a: 25–27). The most recent survey (Hashimoto 2006) cites English donor words at a level of 91% of all loanword tokens. At the same time, the share of other donor languages has plummeted, often drastically (e.g. Dutch and Portuguese to only 1–2%). French donor words currently make up around just 5% of all types (M. Itō 2003: 47; NINJAL 1987: 65) and German a mere 4% of all tokens (Hashimoto 2006).

The oldest English loanword is probably *gere+borotaN* 'Great Britain' (now *gureeto+buriteN*), found in the early 17th century diplomatic chronicle *Ikoku Nikki* (*Chronicle of Foreign Countries*), compiled by the Zen monk Ishin Sūden.[7] The English had a trading post at Hirado (modern Nagasaki prefecture) between 1613 and 1623. English loans did not begin to appear in any number until the mid-19th century, contemporaneous with the appearance of sailors' pidgins in Nagasaki and Yokohama (Atkinson 1879; Daniels 1948; Inoue 2007) and with the arrival in Uraga

---

6 Based on 10 statistical analyses conducted across a range of media since 1980. See Irwin (2011a: 19) for a list (dictionary analyses are ignored).

7 By 1604, James I of England and VI of Scotland was styling himself 'King of Great Brittaine, France and Ireland'. Patterns of phonological adaptation indicate that *gere+borotaN* cannot be derived from a Portuguese, Spanish, or Dutch donor etymon, the only other European donor language candidates of the time.

in 1853 of US Commodore Perry and his Black Ships. An influx of loanwords related to American trends in sport, fashion, and music in the early 20th century saw the donor variety shift from the UK to the US. The post-war US occupation further strengthened this drift, which has since intensified further with globalization and Japanese government education policy. It is no exaggeration to say that any English word or phrase is currently fair game for borrowing, with politicians and bureaucrats frequently criticized for being overly prone to reeling off loanwords for purposes obfuscatory, mystificatory, or abstruse: e.g. *iɴšideɴto* 'incident', *daibaašiti* 'diversity', *suteekuhorudaa* 'stakeholder'.[8]

1.3 Phonological adaptation and phonemicization

This huge influx of English loanwords has meant that traditional "conservative" phonemicizations (Yamaguchi 2007: 16; Tsujimura 2007: 22–37) are no longer tenable. I thus employ the "contemporary" phonemicization of Japanese shown in Figure 1. Similar analyses have been posited by Vance (1987; 2008), Irwin (2011a: 72–74) and Tranter and Kizu (2012). Figure 1 contains 21–23 consonant phonemes,[9] depending on the inclusion of ŋ, whose phonemic status is open to debate (Vance 2008: 214–222), and *v*, generally realized [β], which most speakers have great difficulty pronouncing and found only in highly innovative pronunciations. Other phonemes requiring comment are:

- the voiced velar plosive *g* is typically [g], although may be [ŋ] or [ɣ] for some speakers, chiefly word-internally. See Kindaichi ([1942] 1967); Hibiya (1995, 1996, 1999); Labrune (2012a: 78).
- the affricates *c z č j* are realised [ts dz cɕ ɟʑ].
- the fricatives *f š* are [ɸ ɕ].
- *h* is [ç] before *i y*, but [h] elsewhere.
- the mora obstruent *Q* is the initial segment of a geminate consonant and has a variety of realizations.
- the mora nasal *N* also exhibits a variety of realizations, including [m n ŋ ɲ ɴ]. See Vance (2008: 96–112).

---

8 For greater detail on the history of English donor *gairaigo* than space here can provide, see Umegaki (1963: 71 95); Loveday (1996: 59 76); Yamada (2005: 73 196); Irwin (2011a: 53 61). For a summary of attitudes towards (mainly English) *gairaigo*, see NINJAL (2006); Irwin (2011a: 193 206).

9 Vowel phonemicization is identical in both conservative and contemporary phonemicizations and not shown in Figure 1. The five vowel phonemes /a e i o u/ are broadly realized [ɑ e i o ɯ].

	bilabial		alveolar		(alveolo)-palatal		velar		glottal	moraic
plosive	p	b	t	d			k	g		
affricate			c		č	j				Q
				z						
fricative	f	(v)	s		š				h	
nasal		m		n						N
tap				r						
glide						y		w		

Figure 1: The Japanese consonants: A contemporary phonemicization

Japanese is a mora-timed language, with moras coming in three types: the simple mora *(C)(G)V* (where *G* = an approximant glide *y* or *w*) and the two special moras *Q* and *N*. However, the syllable also has an important role to play (McCawley 1968; Kubozono 1999b).[10] Syllables come in three types: light, heavy, and superheavy. Light syllables consist of a single simple mora *(C)(G)V*. Heavy syllables are bimoraic and consist of a simple mora followed by a vowel or a special mora: thus *(C)(G)VV, (C)(G)VQ* or *(C)(G)VN*. Superheavy syllables are trimoraic and highly marked: *(C)(G)VVQ, (C)(G)VVN* or *(C)(G)VNQ*. Syllable type has a crucial role to play in compound clipping (Section 3.1).[11]

As noted in Section 1.1, during the borrowing process English donor words undergo phonological adaptation, by which donor phones are adapted to the Japanese phonemic inventory. Most phonological adaptation occurs in the form of phonic substitution, but epenthesis (of a vowel or *Q*), as well as deletion, may also take place. Most adaptation pathways have an orthographic source, although auditory sources, while considerably less common, are also found. A detailed description of phonological adaptation is provided in Irwin (2011a: 76–126).

1.4 Compounding

Any of the three vocabulary strata outlined in Section 1.1 can combine with one another during compounding to produce hybrid compounds: e.g. the Sino-Japanese + loan hybrid *deNši+buQku* 'electronic book'. Such hybrids will not be considered in this chapter.

---

10 Some (e.g. Vance 2008: 132) have queried the status of the superheavy syllable in standard (Tokyo) Japanese, while others have refuted the existence of the syllable altogether, most recently Labrune (2012b).

11 Itô and Mester (2015) note an extremely rare fourth syllable type, the 'ultra superheavy'. These come in one tetramoraic form only, *(C)(G)VVNQ*, and must contain a morphological juncture immediately prior to a *Q* initial particle (e.g. *Qte*) or *Q* initial derived suffix (e.g. *Qpoi*). They thus have no role to play in compound clipping.

In the discussion to follow, an important distinction must be drawn between 'imported' and 'assembled' English loanword compounds (Irwin 2011a: 143). I define imported compounds as those that have been borrowed directly from English: e.g. hunger strike → *haNgaa+sutoraiki*. Included among imported compounds are a small number whose elements have been switched during the borrowing process:

(2) a. toaster oven → *oobuN+toosutaa* 'toaster oven'
     b. news flash → *furaQšu+nyuusu* 'news flash'
     c. cabbage roll → *rooru+kyabecu* 'cabbage roll'
     d. off-season → *šiizuN+ofu* 'off-season'

Assembled compounds, meanwhile, are those which have been constructed from two independently borrowed non-compound English loanwords: e.g. resort + lover → *rizooto+rabaa* 'holiday romance'. Assembled compounds are frequently labelled in the literature as *waseieigo* (see Section 1.1). Imported compounds greatly outweigh assembled, although the morphosemantics of the latter are of great interest and explored in Section 2.4. Whether a given compound is imported or assembled is relevant for some truncatory processes (Section 3.2), but irrelevant for others (Sections 3.1 and 3.3).

The morphology of English loanwords in Japanese may be divided into truncatory and non-truncatory phenomena. The latter will be outlined in Section 2. It is the former which are of the greater linguistic interest, however, and analyses of these in Section 3 form the bulk of this chapter.

2 Non-truncatory morphology

After looking at the core morphology of English loanwords in Section 2.1, discussion will move on to affixes and boundedness in Section 2.2, rendaku in Section 2.3 and morphosemantic compounding in Section 2.4.

2.1 Core morphology

The overwhelming majority of English loanwords are nouns[12] and, like all Japanese nouns, appear unmarked for number or case. Any English loan noun may also appear as a verb by means of the native Japanese verbalizer *suru* 'do'. The English

---

12 Some 95% of all loanwords are nouns according to NINJAL (1964). There is no reason to suppose English loanwords behave differently, and little likelihood these figures have changed to any signif icant extent since the survey in question was carried out.

etymon may be a noun, verb, gerund, adjective, preposition or even abbreviation or interjection:[13]

(3) Noun: harassment → *harasumeNto* (*suru*) 'harassment (to harass)'

    Verb: get   → *geQto* (*suru*)   'score, win (to win, score, get)'

    Gerund: heading   → *heQdiNgu* (*suru*)   'header (to head (in soccer))'

    Adjective: live   → *raibu* (*suru*)   'live (to play live)'

    Preposition: off   → *ofu* (*suru*)   'off (to turn off)'

    Abbreviation: Q&A   → *kyuuaNdoee* (*suru*) 'Q&A (to conduct a Q&A)'

    Interjection: bye-bye → *baibai* (*suru*)   'bye-bye (to say bye-bye)'

A very few English loanwords appear as either non-*suru* verbs or as adjectives. These exhibit morphological alternation. As illustrated in (4a) and shown in non-past form, non-*suru* verbs are formed by suffixing -*r* to the English loan, unless this ends in -*ru*, in which case a suffix is unnecessary (4b). These non-*suru* verbs are conjugated without exception as C-final stems: i.e. *wikiQta* 'wikied', never *wikita.[14] Their preferred stem length is 2–3 moras and longer forms are thus frequently clipped (4c), indicated here and in the rest of this section by a double strikethrough (⹀). The verbalization process in (4) is highly marked, however. The vast bulk of English loanword verbs appear as in (3) and many of those that do follow the marked pattern in (4) possess doublets that follow the unmarked pattern in (3).

English loanwords appearing as adjectives do so as either nominal adjectives (5a) or true adjectives (5c), the former being the unmarked pattern. These adjectives undergo the standard morphological processes found in all Japanese adjectives and may also be converted to adverbs via standard morphological rules (5b, 5d).

(4) a. diss   → *disu*   *disu-r-u*   'to diss'

     b. trouble → *toraburu* *torabu-r-u* 'to have trouble'

     c. taxi   → *taku~~šii~~*   *taku-r-u*   'to go by taxi'

(5) a. wild → *wairudo* *wairudo-na* 'wild (out of control)'

     b. wild → *wairudo* *wairudo-ni* 'wildly'

     c. pink → *piNku*   *piNku-i*   'pink, erotic'

     d. pink → *piNku*   *piNku-ku*   'erotically'

---

13 Of course, in many cases the English etymon can belong to more than one part of speech. Here, I have endeavored to use etyma that belong to just one.

14 For theoretical discussion of the *r* suffix, see Chapter 1 (Kageyama and Saito, this volume).

Typically, any inflectional morphology an English etymon may possess has, until fairly recently, been deleted during the borrowing process. Deletion of noun and verb inflectional morphology is illustrated in (6a–e). Derivational morphology may be deleted also, as exemplified in (6f–g), although such deletion has not been as common. More recent borrowings have tended to preserve both the original inflectional and derivational morphology, as shown in (6h–k). This tendency appears to be growing stronger due to an increased awareness among Japanese speakers of English morphology. Indeed, a form of hypercorrection may even occur where an English plural morpheme may be accreted during borrowing (6l–m).[15]

(6)	a.	pyjama**s**	→ *pajama*	'pyjamas'
	b.	wom**en's** lib	→ ***uuma**N+ribu*	'women's lib'
	c.	Rubik**'s** Cube	→ *ruubiꝺku+kyuubu*	'Rubik's Cube'
	d.	corn**ed** beef	→ *kooN+biifu*	'corned beef'
	e.	happy end**ing**	→ *haꝺpii+eNdo*	'happy ending'
	f.	engage**ment** ring	→ *eNgeeji+riNgu*	'engagement ring'
	g.	lucky seven**th**	→ *raꝺkii+sebuN*	'lucky seventh (inning)'
	h.	need**s**	→ *nii**zu***	'needs, demands'
	i.	inform**ed** consent	→ *iNfoomu**do**+koNseNto*	'informed consent'
	j.	idl**ing**	→ *aidori**Ngu***	'(engine) idling'
	k.	empower**ment**	→ *eNpawaa**meNto***	'empowerment'
	l.	top	→ *toꝺpu**su***	'top (garment)'
	m.	bun	→ *baN**zu***	'(hamburger) bun, bap'

Although relatively infrequent, reduplication can occur, usually accompanied by truncation. In the native Japanese and Sino-Japanese strata reduplication typically signals plurality or distribution. With loanwords, however, reduplication appears to function in a manner similar to the mimetic stratum, signalling emphasis or intensity:

(7)	a.	love	→ *rabu+rabu*	→ *raburabu*	'lovey-dovey'
	b.	erotic	→ *ero~~ꭒꝺku~~+ero~~ꭒꝺku~~*	→ *eroero*	'porno'
	c.	grotesque	→ *guro~~tesuku~~+guro~~tesuku~~*	→ *guroguro*	'gross'

---

15 A hypercorrection in the sense that the plural of a countable English noun occurs with vastly greater frequency than the singular. The examples given ignore donor nouns which typically occur in pairs or only rarely in the singular.

In their verbalization and adjectivization strategies (3–5), the loanword and Sino-Japanese strata – both borrowed, albeit at greatly different time depths – behave similarly. More divergent strategies are exhibited by the mimetic stratum: its verbalization strategies are the same as both the loanword and Sino-Japanese strata, while its adjectivization strategies are more idiosyncratic. The native stratum is the outlier: its *r*-suffix, is "semi-productive" (Itô and Mester 2015), while its adjectivization strategies are, broadly, the opposite of the loanword and Sino-Japanese strata.

Verbalization and adjectivization strategies across all four vocabulary strata are summarized in Table 1, along with examples. Markedness is indicated by means of a crude 4–point scale, where 3 = 'unmarked', 2 = 'somewhat marked', 1 = 'highly marked', and 0 = 'non-existent'.

Table 1: Verbalization and adjectivization strategies across vocabulary strata

	LOAN	SINO-JAPANESE	NATIVE	MIMETIC
suru verbalization	3	3 *kokuhaku suru* 'confess'	3 *tamuro suru* 'hang out'	3 *jimejime suru* 'be damp'
r-suffix	1	1 *koku r u* 'cough up, fess up'	2 *futo r u* 'get fat'	1 *paku r u* 'lift, crib'
nominal adjectives	3	3 *daiji na* 'important'	1 *čiisa na* 'small'	2 *gunyagunya na* 'mushy'
true adjectives	1	1 *šikaku i* 'square'	3 *čiisa i* 'small'	0

2.2 Affixes and boundedness

Table 2 offers some examples of English loanwords, frequently prepositions or affixes, functioning as prefixes or, less commonly, suffixes in Japanese. There are no examples of infixes. In most cases, semantic change has occurred, as is clear from the English gloss. In most cases also, the Japanese affixes are bound. Those few that may also occur unbound frequently show a semantic difference between their bound and unbound form (e.g. bound *širubaa-* 'senior citizen, OAP', but unbound *širubaa* 'silver'). A few affixes have not yet reached a sufficient level of productivity for inclusion in Table 2. These include the suffixes *šiQpu* ← -ship (e.g. skin → *sukiN+šiQpu* 'close physical contact, esp. between mother and child') and *-fuuru* ← -ful (e.g. heart → *haato+fuuru* 'heart-warming').

Table 2: English loanwords as affixes: A selection

<table>
<tr><td></td><td>AFFIX</td><td>GLOSS</td><td colspan="2">EXAMPLES</td></tr>
<tr><td rowspan="5">PREFIXES</td><td>my → mai</td><td>'personal, one's own'</td><td>mai+kaa ← car
mai+hoomu ← home
mai+buumu ← boom</td><td>'private car'
'owner occupied house'
'personal fad'</td></tr>
<tr><td>no → noo</td><td>'un , not'</td><td>noo+taQči ← touch

noo+kauNto ← count
noo+kaQto ← cut</td><td>'steering clear (of a potential problem)'<br>'not counted, not included'
'uncut (movie)'</td></tr>
<tr><td>the → za</td><td>'the only, the best'</td><td>za+terebijoN ← television
za+šiNfonii+hooru ← symphony hall
za+šinema ← cinema</td><td>TV website
Osaka concert hall

satellite movie channel</td></tr>
<tr><td>silver → širubaa</td><td>'senior citizen, OAP'</td><td>širubaa+šiito ← seat
širubaa+saabisu ← service
širubaa+maNšoN ← mansion</td><td>'priority seat'
'care for the elderly'
'sheltered housing'</td></tr>
<tr><td>over → oobaa</td><td>'surfeit, complete'</td><td>oobaa+doraQgu ← drug
oobaa+dokutaa ← doctor
oobaa+neQto ← net</td><td>'overdose'
'overterm PhD candidate'
'reaching over the net (volleyball)'</td></tr>
<tr><td rowspan="4">SUFFIXES</td><td>off → ofu</td><td>'limit, deduction'</td><td>price → puraisu+ofu
book → buQku+ofu

pay → pee+ofu</td><td>'discount'
second hand book store chain
'financial cap'</td></tr>
<tr><td>man → maN</td><td>'person'</td><td>gag → gyagu+maN
guard → gaado+maN
salary → sararii+maN</td><td>'comedian'
'security guard'
'white collar worker'</td></tr>
<tr><td>er → raa</td><td>'obsession'</td><td>ketchup → kečaQpu+raa
Uniqlo → yunikuro+raa
Wiki → wiki+raa</td><td>'ketchupaholic'
'Uniqlo freak'
'Wikipedia nut'</td></tr>
<tr><td>down → dauN</td><td>'reduction'</td><td>image → imeeji+dauN
speed → supiido+dauN
price → puraisu+dauN</td><td>'damage to reputation'
'slowdown'
'reduction in price'</td></tr>
</table>

The prefix *mai-* is not restricted to the first person, but can be used with reference to the second or third person also. The suffix *-raa* is a reanalysis based on the large number of English donor words ending in <-ler, -rer, -lor > → *-raa*: e.g. Internet Explo**rer** → *intaaneQto+ekusupuroo**raa***, counse**llor** → *kauNse**raa***, hust**ler** → *hasu**raa***, etc. Truncation is not uncommon with this suffix. A reanalysis similar to *-raa* can be found with two other suffixes not listed in Table 2: *-tariaN* ← <(t)arian> and *-čiQku* ← <(t)ic>. These suffixes rarely combine with English loans however, favoring instead non-loan morphemes: e.g. *jibe+tariaN* 'young person who squats on the ground with others' or *maNga+čiQku* 'manga-esque'.

The loanword stratum is not unique in providing affixes, and their behavior parallels very closely that found in other strata: most are bound and infixes do not occur at all (see Chapter 1 [Kageyama and Saito, this volume] for further detail).

Although not occurring as affixes, some English loanwords occur extremely rarely, or never, unbound. *Kaa* ← car and *sukuuru* ← school, for example, appear with overwhelming frequency as bound morphemes in loanword compounds: e.g. *kaa+diiraa* ← car dealer, *patorooru+kaa* ← patrol car (= 'police car'), *sukuuru+zooN* ← school zone, *samaa+sukuuru* ← summer school. When an unbound morpheme is required, the non-loans *kuruma* / *jidooša* 'car' and *gaQkoo* 'school' are preferred.

2.3 *Rendaku*

Perhaps the mostly widely researched phenomenon observed in Japanese compounding is *rendaku*, or sequential voicing, where the initial voiceless obstruent of a non-initial element may undergo voicing (Vance 2015). However, it has been well established (Nakagawa 1966; Kubozono 1999a; Takayama 2005; Tamaoka et al. 2009; Irwin 2011a: 150–153) that *rendaku* does not occur to any extent in the loanword stratum. Here, it is restricted to only a very few Portuguese loans borrowed in the 17th century and a sole compound containing a truncated English loanword, the now obsolete *akageQto* 'country bumpkin' (lit. 'red blanket'), where *keQto* is the mora-clipped (see (14) below) output of *buraNkeQto* ← blanket.

2.4 Morphosemantic compounding

Many assembled compounds exhibit severe semantic divergence from the English donor elements from which they are composed. I will term this phenomenon 'morphosemantic compounding'. When encountered by English-speaking, or English-conversant, learners of Japanese, morphosemantic compounds are frequently parsed incorrectly. As Daulton (2008: 19) has noted, this, combined with the fact that Japanese learners of English are frequently inclined to presume morphosemantic compounds are correct English (i.e. imported rather than assembled), leads to difficulties with foreign language learning in both directions.

I do not include in my definition of morphosemantic compounds those assembled compounds containing an English loanword functioning as an affix (Section 2.2). English loanwords functioning as affixes have attained semantic stability and such stability generally precludes incorrect parsing: an English-speaking, or English-conversant, learner of Japanese has become aware of the meaning of such affixes through having encountered them frequently across a range of compounds.

With a few exceptions, morphosemantic compounds are restricted to five broadly definable semantic domains: marketing and media (8), technology (9), automotive

(10), sports (11), and sex (12). The motivation behind morphosemantic compounding in these domains is not hard to ascertain. With the first three – media, technology, and automotive – the motive is prestige. English loans are perceived by Japanese speakers as having high status. TV and radio programs, websites, advertisements, high-tech merchandise, and high-end consumer goods, such as automobiles, are heavy users. Need also plays a role: many new technologies and media have no non-loan Japanese equivalents.[16] The sheer number of English loans which can be encountered in these three domains (JKS (ed.) 2013) provides ample reproductive fodder for morphosemantic compounding. In the domain of sports, the motive is perhaps best articulated as 'power'. The prestige accorded the English language and English loans confers a sense of power on morphosemantic coinages, and power is a *sine qua non* of the sporting world.[17] Finally, the prime motivation with the domain of sex (as well as the sex industry and sexual bodily functions) is euphemism. A similar trend is found in English, where Latin and French borrowings are especially prevalent. Here, morphosemantic compounds can contain elements of otherwise exceptionally low frequency to further obfuscate meaning and heighten euphemism. The few morphosemantic compounds which fall outside these five domains (8)–(12) also tend towards euphemistic usage and include green + car → *guriiN+kaa* 'first-class railway carriage', ceremony + hall → *seremonii+hooru* 'funeral hall' and hello + work → *haroo+waaku* 'government job seekers bureau'.

(8) time + service → *taimu+saabisu* 'special offer, special sale'
    golden + hour → *goorudeN+awaa* 'prime time TV'
    mic + location → *maiku+rokeešoN* 'on-the-spot broadcasting'

(9) catch + phone → *kyaQči+hoN* 'call waiting'
    plus + screwdriver → *purasu+sukuryuudoraibaa* 'Phillips screwdriver'
    ten + key → *teN+kii* 'numeric keypad'

(10) gasoline + stand → *gasoriN+sutaNdo* 'petrol station, gas station'
     front + glass → *furoNto+garasu* 'windscreen, windshield'
     door + engine → *doa+eNjiN* 'auto open-and-close (taxi) door'

---

16 Within the automobile domain, another possible motivation may be donor confusion. Many car parts, as well as automobile related goods and accessories, differ markedly across English varieties: the gulf between the UK and the US in this respect is huge (Hargraves 2003: 179 187). It can be unclear to Japanese speakers what the 'correct' English term actually is and morphosemantic compounding can thus be seen as a fudge.

17 Note that in another aspect of sport, fair play, English loans are rarely encountered. The Japanese notion of what constitutes gentlemanly conduct or sportsmanship does not require English loan words for adequate expression.

(11) out + course → *auto+koosu* 'front nine (golf)'
   toss + batting → *tosu+baQtiNgu* 'pepper drill (baseball)'
   point + getter → *poiNto+geQtaa* 'goal scorer (soccer)'

(12) love + hotel → *rabu+hoteru* 'hotel renting rooms by the hour'
   delivery + health → *deribarii+herusu* 'prostitute call-out service'
   blue + day → *buruu+dee* 'day when one is menstruating'

Many of the morphosemantic compounds illustrated in (8)–(12) undergo a range of truncatory processes and it is to these that I turn in the following section.

3 Truncatory morphology

Three truncatory processes occur with English loanword compounds: compound clipping (Section 3.1), morpho-orthographic truncation (Section 3.2), and ellipsis (Section 3.3). These are illustrated in Table 3 by means of the imported loan compound *dejitaru+kamera* ← digital camera. The first two of these processes, compound clipping and morpho-orthographic truncation, are morphophonological phenomena, while the last, ellipsis, is purely morphological.

Table 3: Truncation types

TRUNCATION TYPE	INPUT	OUTPUT
compound clipping (Section 3.1)	**deji**taru+**kame**ra	*dejikame*
	dejitaru+**ka**mera	*dejika*
	dejitaru+**kame**ra	*dekame*
	dejitaru+ka**mera**	*dejimera* ...... etc.
morpho-orthographic truncation (Section 3.2)	**d**igital + **c**amera	
	↓ ↓	*DC*
	dejitaru+kamera	
ellipsis (Section 3.3)	**dejitaru**+kamera	*dejitaru*
	dejitaru+**kamera**	*kamera* ...... etc.

If a compound undergoes truncation via one of the three processes illustrated, it typically does so via one and one only. If multiple truncated forms do occur, then usage is skewed towards one of them. The unmarked word for 'digital camera' in Japanese is *dejikame*, via the first of the truncatory options in Table 3. This clipped form is encountered far more frequently than its full form *dejitaru+kamera*. Indeed, of the three truncation processes to be examined, compound clipping is by far the most frequent. However, other truncations of *dejitaru+kamera* may be found, typically in youth speech, jargon, slang, and dialect, indicating that truncation processes possess a strong sociolinguistic component. The other, highly marked, trunca-

tions of *dejitaru+kamera* shown in Table 3 are only some of the options available: some of those shown may not actually be encountered, while others that may be encountered are not shown. Truncation processes are extremely fluid and the truncations themselves frequently highly transient.

All three types of truncation shown may be hybrid, i.e. the compound input consists of a mixture of English and non-English loanwords (e.g. ice + Dutch *koffie* 'coffee' → ***ai*su+*ko*ohii** → *aiko* 'iced coffee') or a mixture of English loanword and non-loanword (e.g. service + Sino-Japanese *zaNgyoo* 'overtime' → ***saabi*su+*zaN*gyoo** → *sabizaN* 'unpaid overtime'). Such hybrids will not be treated here.

For truncatory processes to operate implies an awareness on the part of the Japanese native speakers applying the process that the loanword to be truncated is a compound, despite this often being opaque in the case of imported compounds. Such an awareness is the only explanation for the fact that, in compound clipping (Section 3.1), moras are clipped either from the beginning or from the end of individual elements (never from the middle, never at random) and that, in ellipsis (Section 3.3), an entire element is deleted. For morpho-orthographic truncation (Section 3.2) we must presume, in addition, knowledge of the original English spelling.

3.1 Compound clipping

All examples cited in this section occur in the author's English Loanword Clipped-Compound Integrated Database (hereafter "El Cid"). All El Cid compounds (*n* = 432) contain two elements, both of which have been borrowed from English and appear as either an entry (or as a redirection to an entry) in the Wikipedia Japan pages posted since 2009 or in one of the following sources: Umegaki (1956, 1963); Arakawa (ed.) (1977); JKS (ed.) (1977, 2003, 2010); Maruyama et al. (1992); Yonekawa (1997); GJH (eds.) (2001); Watanabe, Skrzypczak and Snowden (2003); Shinmura (ed.) (2008); SSDH (2010).[18] As a database, I make no claims for El Cid being 100% comprehensive. That said, I believe it to be an accurate reflection of compound clipping realities with respect to the statistical trends to be discussed below.

In previous studies a variety of different names has been employed for what I will term compound clipping. J. Itô (1990: 220) refers to the phenomenon as 'compound abbreviation', while Labrune (2002: 102–104) calls the output of the process 'compound abbreviated loanwords'. This fluid nomenclature reflects the general lack of agreement on a formal subdivision for the truncatory morphology of loanwords.

---

18 Compounds consisting of a truncated reduplicated loanword, as in (7), are not included in El Cid. However, compounds containing elements which may have been borrowed from either English or French are included: e.g. *omurecu* 'omelette', *restoraN* 'restaurant'. See Irwin (2011a: 50) for discussion.

Restricting discussion in line with El Cid content to compounds containing two elements, I define compound clipping as:

(13)   the reduction of a loanword compound to between 2 and 6 moras (μ), by means of clipping at least one mora from at least one element. Moras are clipped overwhelmingly from the rear of an element and the process applies equally to both imported and assembled compounds.

There exists a range of possible outputs and these are expressed below as, for example, 1μ+1μ. In this case, the output is a 2μ truncation composed of two elements both of which have been clipped back to one mora only. Compound clipping is illustrated in Table 4.

Compound clipping should not be confused with mora-clipping. The latter is a purely phonological phenomenon, as illustrated in (14a): see Irwin (2011a: 131–136) or Kubozono (2015) for further examples. In mora-clipping the syllable plays a major role (J. Itô 1990; Itô and Mester 1992; Labrune 2002), whereas in compound clipping, as we shall see below, the syllable plays only a bit part. With mora-clipping, suprasegmental factors (i.e. accent) are critical in determining where a loanword will be clipped (Labrune 2002; Irwin 2011b); with compound clipping, suprasegmental factors appear to be irrelevant. This does not mean, however, that English loanword compounds may not undergo mora-clipping. Although such cases are rare, the examples in (14b) testify to their existence. Such truncation falls outside my definition of compound clipping in (13) and will not be treated further.

(14)   a.   chocolate   →   ***čoko**reeto*   →   *čoko*
           television   →   ***terebi**joN*   →   *terebi*

       b.   convenience store →   ***koNbini**eNsu+sutoa* →   *koNbini*
           concentric plug   →   ***koNseNto**riQku+puragu* →   *koNseNto*

It is clear from Table 4 that the unmarked clipping process is truncation to a 2μ+2μ type: such types comprise 72% (313/432) of El Cid. The only other type to occur to any significant extent is the 2μ+1μ type, which makes up 16% (67/432) of the database. Other types are extremely rare and occasionally obsolete: e.g. the two 1μ+1μ types shown in Table 4, *moga* and *bea*. That compound clippings were heavily skewed towards 2μ+2μ types was pointed out, without the benefit of a database, as long ago as Umegaki (1963: 108–109).

In the vast majority of cases, the clipped elements of which clipped compounds are composed do not exist as unbound morphemes. In other words, *pi* and *bairiN* are

Table 4: Examples of compound clipping

TYPE	DONOR ENGLISH	JAPANESE INPUT	TRUNCATED OUTPUT	ENGLISH GLOSS	EL CID %
1μ+1μ	modern + girl	**mo**daaɴ+**ga**aru	moga	flapper	1.2%
	base + up	**be**esu+**a**Qpu	bea	basic pay increase	
1μ+2μ	light novel	**rai**to+**nobe**ru	ranobe	light novel	2.1%
	mail address	**me**eru+**ado**resu	meado	email address	
1μ+3μ	picture + crossword	**pi**kučaa+**kurosu**waado	pikurosu	nonogram (logic puzzle)	1.4%
	Pacific + league	**pa**šifiQku+**riigu**	pariigu	Pacific (baseball) league	
2μ+1μ	Brad Pitt	**bura**Qdo+**pi**Qto	burapi	Brad Pitt	15.5%
	smart phone	**suma**ato+**ho**ɴ[19]	sumaho	smart phone	
2μ+2μ	unique + clothing	**yuni**iku+**kuro**ojiɴgu	yunikuro	Uniqlo	72.5%
	costume + play	**kosu**čuumu+**pure**e	kosupure	cosplay	
2μ+3μ	cutlet + sandwich	**kacu**recu+**saɴdo**iQči	kacusaɴdo	wiener schnitzel sandwich	4.2%
	husband + hunt	**hazu**baɴdo+**haɴto**	hazuhaɴto	husband hunting	
2μ+4μ	self + motor	**seru**fu+**mootaa**	serumootaa	self starting motor	0.7%
3μ+1μ	black + lagoon	**bura**Qku+**ra**guuɴ	burakura	Black Lagoon (computer game)	0.5%
3μ+2μ	cool + business	**kuuru**+**bizu**nesu	kuurubizu	dressing down for the heat	1.2%
	auto + bicycle	**ooto**+**bai**šikuru	ootobai	motorbike	
3μ+3μ	plastic tile	**purasu**čiQku+**tairu**	purasutairu	plastic tile	0.5%
4μ+2μ	bilingual + girl	**bairiɴ**garu+**gyaru**	bairiɴgyaru	young female returnee	0.5%

found only bound in compounds, never as unbound mora-clipped forms (14) of the loans *pikučaa* 'picture' and *bairiɴgaru* 'bilingual'.[20]

The most favored total length for clipped output is 4μ (74%, 321/432), although 3μ output is not uncommon (18%, 76/432). Both initial (401/432, 93%) and final elements (329/432, 76%) are clipped most often to 2μ, this proclivity being far stronger among initial elements. In terms of element balance, nearly three-quarters of El Cid compound clippings have both elements clipped to an identical length. As far as the remaining quarter is concerned, a longer first element (76/432, 18%) is favored over a longer final element (36/432, 8%). All these trends are summarized in Figure 2.

The dominant 2μ+2μ output witnessed for compound clipping, as well as the fact that both initial and final elements are most commonly clipped to 2μ, are all yet further examples of the importance of the foot (=2μ) in Japanese phonology, morphology, and morphophonology. Not only is the foot the most frequently found length for native Japanese nouns (see Itô and Mester 2015), it is also the preferred length for loanword back-clippings (Irwin 2011b; see (14a) for an example). It also plays an important role in non-loan compound clipping (Shibatani 1990: 254–256),

---

19 Also, *sumaato+foɴ*.

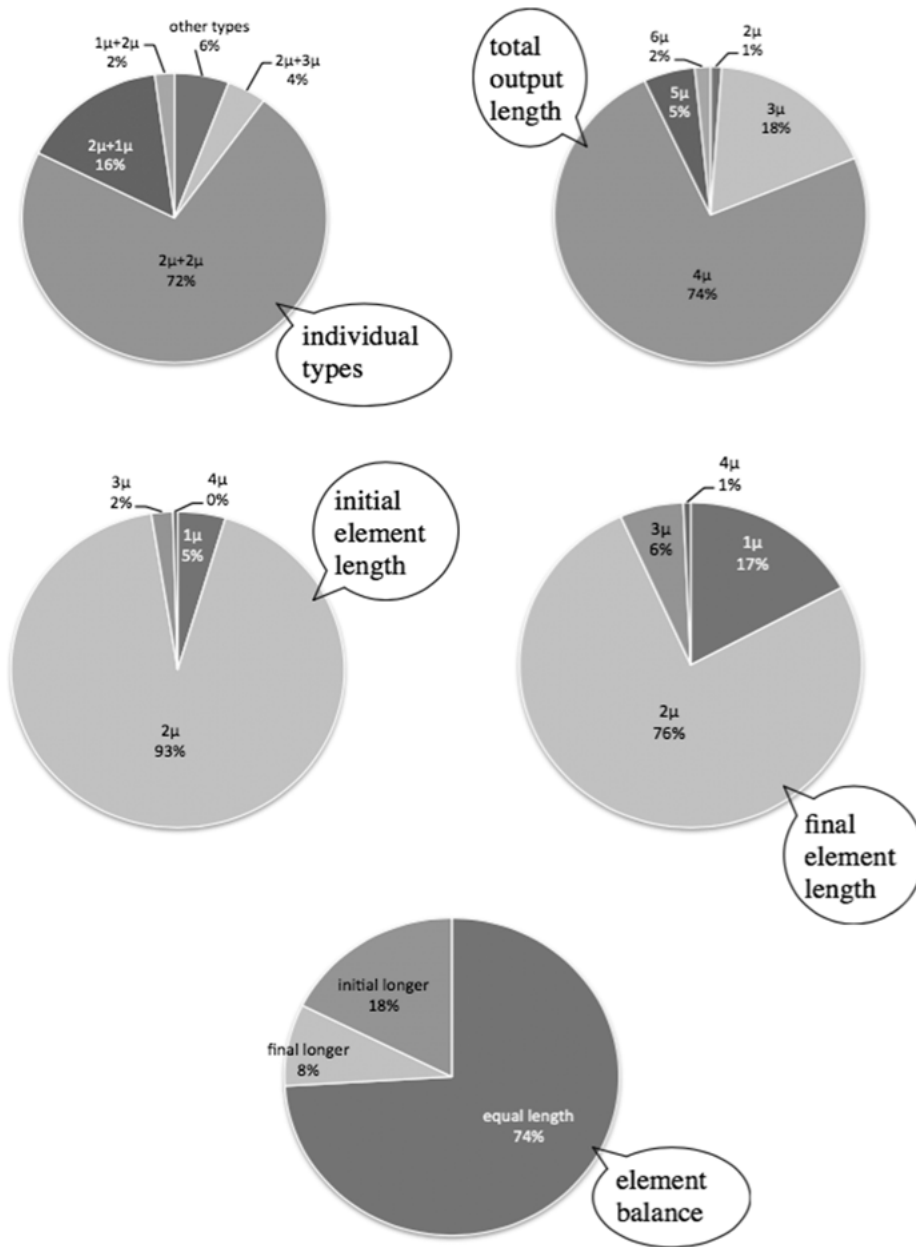**20** The small number of exceptions include *kacu* 'cutlet' and *saɴdo* 'sandwich'.

Figure 2: Analysis of compound clipping output

in the accentuation patterns found in compounds (Tsujimura and Davis 1987), in rendaku immunity (Rosen 2003; Irwin 2009a), in mimetic reduplication (Poser 1990: 94–95; Hamano 1998: 25–38), in the sociolinguistic phenomenon of jazz argot (Tateishi 1989; Itô, Kitagawa and Mester 1996), and in hypocoristic formation (Mester 1990; Poser 1990: 81–93).

The question remains, however, as to why approximately one-sixth of all clipped compounds are of the 2μ+1μ type. Clearly, the importance of the foot means a similar proportion of 3μ+1μ or 4μ+3μ types would be unexpected. But why 2μ+1μ in particular? Why not 1μ+2μ or 2μ+3μ?

One possible motivation (Kubozono 2002: 94) is that, in Japanese in general, long vowels (V_aV_a) tend to be shortened in word-final position. Thus, in 2μ+1μ type clipped compounds, the final element is clipped to 1μ in order to avoid a 2μ CV_aV_a heavy syllable. Had they been clipped to the dominant 2μ+2μ output, 79% (53/67) of 2μ+1μ types in El Cid would have had a final element consisting of a single heavy syllable (Section 1.3), and 60% (40/67) a final element consisting of a CV_aV_a heavy syllable (see Figure 3). On the other hand, a similar analysis of the actual 2μ+2μ output found in El Cid raises difficulties in maintaining the argument that word-final long vowels are disfavored. Here, the proportions were smaller: 40% (124/313) had a final element consisting of a single heavy syllable and 8% (26/313) a final element consisting of a CV_aV_a heavy syllable. Further, as shown in (15), there exist final elements in the Japanese input which may be clipped to either 1μ or 2μ, the 2μ variant containing a final CV_aV_a heavy syllable. Here, neither the syllable structure, phonotactics, nor suprasegmentals of the initial element appear to play any role in the choice of variant.

(15) a. *geemu* ← game
    2μ+1μ ***saba**ibaru+**ge**emu* → *sabage* 'survival game'
    2μ+1μ ***ne***Q*to+**ge**emu*[21] → *netoge* 'internet game'
    2μ+2μ ***reto**ro+**gee**mu* → *retogee* 'retro computer game'
    2μ+2μ ***moba**iru+**gee**mu* → *mobagee* 'mobile phone game'

   b. *boodo* ← board
    2μ+1μ ***poke***Q*to+**bo**odo* → *pokebo* 'add-on keyboard for mobile'
    2μ+1μ ***suno**o+**bo**odo* → *sunobo* 'snowboard'
    2μ+2μ ***suno**o+**boo**do* → *sunoboo* 'snowboard'
    2μ+2μ ***suke**eto+**boo**do* → *sukeboo* 'skateboard'

Another possible motivation, avoidance of a word-final mora obstruent, generally illicit in Japanese, must be discounted. As well as accounting for only a small number

---

21 See **DELETION OF THE MORA OBSTRUENT Q** below for an explanation of the Q deletion found in this clipped compound.

(16%, 11/67) of 2μ+1μ types, there exist alternative strategies for word-final mora obstruent avoidance which preserve an unmarked 2μ+2μ output (see DELETION OF THE MORA OBSTRUENT Q below).

It would thus appear there are no transparent constraints or rules that can account for a dominant 2μ+2μ type clipped compound further truncating its second element to yield a 2μ+1μ type. There is no obvious explanation as to why 2μ+1μ types should be the second most favored compound clipping type. It is likely we are witnessing sociolinguistic variation based on some as yet unknown variable or variables. Here, and with other truncatory morphology to be discussed below, it is hoped future sociolinguistic research can shed greater light on a confused situation.
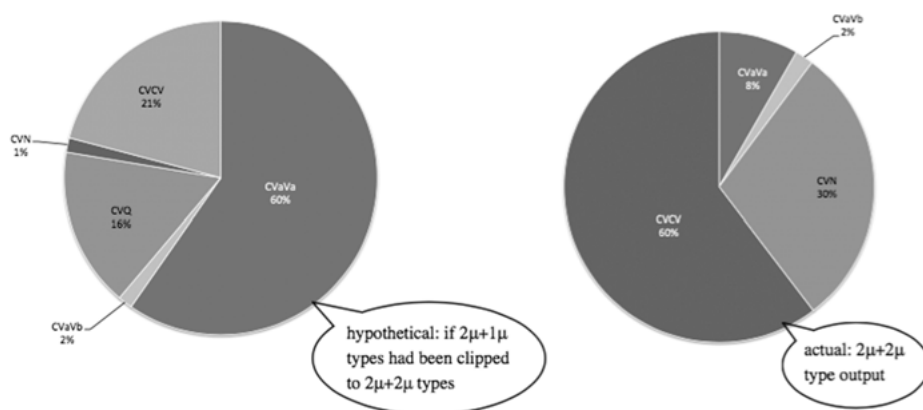

Figure 3: Final element syllable types (hypothetical and actual)

Several further issues related to compound clipping require further discussion:

HALF-CLIPPINGS: Truncations where only one element is clipped – recall my definition of compound clipping in (13) – are termed 'half-clippings' (Irwin 2011a: 144). These include *pariigu*, *hazuhaNto*, *serumootaa*, *kuurubizu*, *ootobai*, *purasutairu* and *bairiNgyaru* (the unclipped element is underlined) in Table 4. Half-clippings are not uncommon and comprise some 18% (77/432) of El Cid.

The motivation behind half-clipping, in many cases, appears to be preservation of the foot. In fully three-quarters of all half-clippings (75%, 58/77) the unclipped element is 2μ, this despite the fact that 1μ-elements are perfectly licit (1μ-elements make up 11% (94/864) of all elements in El Cid). The strong trend towards preserving the foot is even clearer when viewed from another angle: of all the 2μ-elements appearing in the El Cid Japanese loan input, only 6% (4/65) were clipped to 1μ.

In other cases, the motivation behind half-clipping is obscure. In the remaining one quarter of half-clippings, the unclipped element is typically 3μ (23%, 18/77) or,

rarely, 4μ (1%, 1/77) in length.[22] Although in the majority of these cases (68%, 13/19) clipping to 2μ would yield a heavy syllable, unlike mora-clipping (14a) there is no constraint against clipped elements consisting of a single heavy syllable in compound clipping. Thus, any argument that half-clipping finds its motivation in the avoidance of a single heavy syllable must be rejected: there is no reason why we should not find *hazuhaN, *serumoo, *kuubizu or *oobai for hazuhaNto, serumootaa, kuurubizu, and ootobai. Indeed, being unmarked 2μ+2μ truncations, we should in fact expect *hazuhaN, *serumoo, *kuubizu and *oobai as the optimal output. That this does not happen is curious.

NON-FINAL VOWEL DELETION IN NON-LIGHT SYLLABLES: Some 4% (19/432) of El Cid clipped compounds undergo some form of non-final vowel deletion in addition to truncation. Such deletion occurs only within heavy or super-heavy syllables (i.e. non-light syllables) and comes in two types.

The more common of the two (79%, 15/19) is deletion of the second vowel in an element-initial heavy syllable, typically (13/15) in an initial element: e.g. birth control → **baasu+koNtorooru** → basukoN; personal computer → **paasonaru+koNpyuutaa** → pasokoN; sailor pants → **seeraa+paNcu** → serapaN, etc. In all cases, this deleted second vowel is identical to that preceding it, i.e. it cannot be claimed that the heavy syllable contains a diphthong. This type of deletion may thus be termed 'vowel shortening'. Interestingly, initial-element vowel shortening occurs solely with imported compounds, while the only two examples of final-element vowel shortening occur with assembled compounds: hello + morning → **haroo+mooniNgu** → haromoni (name of TV program); and bloomers + sailor (suit school uniforms) → **buruumaa+seeraa** → burusera (shop selling school girl's uniforms and underwear complete with photo ID of wearer). There are no examples of vowel shortening occurring in both elements of a clipping.

The less common type of non-final vowel deletion (21%, 4/19) occurs only in super-heavy syllables ending in -auN. Here, rather than the syllable-final N, it is the preceding u which is deleted (Kuwamoto 1998a; Kubozono 2001a). There are only four examples in El Cid, including soundtrack → **sauNdo+toraQku** → saNtora and Blue Mountain (coffee brand) → **buruumauNteN** → burumaN. This type of deletion occurs with equal frequency in both initial and final elements and across both imported and assembled compounds.

The motivation behind the former type of non-final vowel deletion, vowel shortening, is unclear. On the one hand, all the elements in which vowel deletion operates are reduced to 2μ, thus achieving the optimal element length of one foot. On the other hand, however, there are examples of the second vowel in a non-final heavy syllable not undergoing deletion, in spite of the fact that such deletion would

---

22 There are no unclipped elements (i.e. no El Cid Japanese loan input) 1μ in length. This is due to the phonological adaptation pathways which apply during the borrowing process (Section 1.3) and which mitigate in general against the creation of 1μ English loans. There exists only one, za ← the, which I treat as an affix (Section 2.2).

yield an unmarked 2μ+2μ output: e.g. verification card → ***beri**fikeešoN+**kaado*** → *berikaado*, as well as *kuurubizu* and *ootobai* in Table 3.

With the latter type of non-final vowel deletion, that of *u* in super-heavy syllables ending in *-auN*, the motivation is clearly one found across the Japanese language in general, that of avoiding super-heavy syllables. Of course, such a syllable could just as easily be avoided by deleting the syllable-final mora nasal *N*. Note, however, an identical type of non-final vowel deletion can occur during the phonological adaptation of English loanwords (Ichikawa 1929, 1930; Lovins 1973: 81–82; Kubozono 1995, 1999b; Irwin 2011a: 100–101): e.g. (sports) ground → *guraNdo*; (cosmetic) foundation → ***faN**deešoN*.

DELETION OF THE MORA OBSTRUENT Q: A very small number (1%, 5/432) of clipped compounds evince *Q*-deletion. Since, in Japanese in general, word-final *Q* is licit only in the mimetic stratum and in emphatic forms, it is unsurprising that the mora obstruent *Q* is deleted in cases where it would become a word-final mora. The two cases where this occurs (16a) yield 2μ+2μ type truncations. However, deletion is not the only strategy found for avoiding illicit word-final *Q*. Clipping the final element to yield a 2μ+1μ output is also possible: see the examples in (16b). Why some Japanese input follows the truncation pattern in (16a) while other input follows the pattern in (16b) is unclear.[23] Once again, it is possible we are dealing with an unknown sociolinguistic variable.

(16) a. 2μ+2μ
      American football → ***ame**rikaN+**fu**Qtobooru* → *amefuto* (*amefuQ)
      Harry Potter → ***hari**i+**po**Qtaa* → *haripota* (*haripoQ)

   b. 2μ+1μ
      Photoshop → ***foto**+**šo**Qpu* → *fotošo* (*fotošoQ, *fotošopu)
      potato chips → ***pote**to+**či**Qpusu* → *poteči* (*potečiQ, *potečipu)

The other three cases of *Q*-deletion in El Cid occur word-internally, despite the fact that word-internal *Q* is perfectly licit in Japanese in general. Nevertheless, the motivation behind its deletion is transparent. Were they to be clipped to the dominant 2μ+2μ type output, the examples in (17a) show that retention of initial-element-final *Q* would result in either an illicit *Qm*, a quasi-illicit *Qr*, or a marked *Qg* gemination.[24]

---

23 A further potential strategy is *Q* → *cu* (Kuwamoto 1998b) under the influence of *kana* orthography, where *Q* つ is written as miniscule *cu* つ. Although there are no examples in El Cid, an example of a clipped hybrid compound where this strategy is apparent is: Sino Japanese *daNzeN* 'absolute' + top → ***daN**zeN+**to**Qpu* (だんぜん + ト ツ プ) → *daNtocu* (だん ト ツ) 'decisive, runaway'.

24 Geminate nasals are phonemically *Nm* and *Nn*. The existence of a geminate liquid *Qr* is moot (Irwin 2011a: 73; Labrune 2012a: 136; Tranter and Kizu 2012: 270), hence my use of 'quasi illicit'. Geminate voiced obstruents (*Qb*, *Qd*, *Qg*) are restricted to the loanword stratum and thus marked in standard Japanese. Even here, they may be subject to devoicing (Arisaka 1940: 94; Quackenbush 1989; Vance 2008: 108 110; Kawahara 2011a, 2011b, 2012).

Evidence for this motivation is further strengthened by the examples in (17b), where word-internal Q is not deleted and the resulting clipped compounds contain licit, unmarked geminations. Disfavored geminations such as those in (17a) may, nevertheless, be dealt with elsewhere not by deletion, but by further clipping the initial element to yield a 1μ+2μ output, as seen in (17c). This is despite the fact that, in some cases, an unmarked 2μ+2μ output would contain an unmarked gemination, such as Qs.

(17) a. (inter)net radio   → ***neQto***+***raji***o   → *netoraji* (*neQraji)
     (browser) bookmark → ***buQku***+***maa***ku   → *bukuma* (*buQma)
     (inter)net game   → ***neQto***+***gee***mu   → *netoge* (*neQge)

    b. hit parade   → ***hiQ***to+***paree***do   → *hiQpare*
     shocking pink   → ***šoQki***ngu+***piN***ku   → *šoQpiN*
     mid-century   → ***miQ***do+***seN***šarii   → *miQseN* 'designer furniture'

    c. Netscape   → ***neQto***+***suke***epu   → *nesuke* (*neQsuke)
     sex + friend   → ***seQ***kusu+***fureN***do → *sefure* (*seQfure)

BLENDS: Here, the front rather than the rear – recall the definition in (13) – of one element is clipped by at least one mora. Blends are infrequent, occurring in 6% (24/432) of clipped compounds in El Cid. In all these cases bar one, it is the front of the final element which is clipped.[25] Further, in the vast majority of cases (88%, 21/24) the Japanese input is an assembled compound. No particular output is dominant. The initial element is clipped to 2μ in only 71% (17/24) of blends and the final element to 2μ in only 54% (13/24) of cases: compare the data for El Cid as a whole in Figure 2. Overall, only 42% (10/24) of blends are of the 2μ+2μ type.

Some examples of blends are shown in (18). Here, the assembled compound (18a) means 'an apartment with a garage', (18c) 'a type of marathon run while carrying food' and (18d) 'a Japanese working in Asia'.

(18)   a.   1μ+3μ
       motor + apartment   →   ***mo***otaa+*a****paato***   →   *mopaato*

     b.   2μ+2μ
       child idol   →   ***čai***rudo+ai***doru***   →   *čaidoru*

     c.   2μ+3μ
       marathon + picnic   →   ***mara***soN+piQku***niQku*** →   *maraniQku*

     d.   3μ+2μ
       Japanese + Asian   →   ***japani***izu+aji***aN***   →   *japaniaN*

---

25 The sole exception is model + casual → *mo****deru***+***kaju***aru→*derukaji* 'casualwear as worn by models'. This blend also undergoes **MORA SPLITTING**.

MORA SPLITTING: Mora-splitting may occur only where the Japanese input contains a kana digraph, i.e. a single mora written using two *kana* (see Section 3.2), the second of which is written in miniscule: ジャ <ja>, ティ <ti>, フェ <fe>, etc. The mora in question is split by clipping the second miniscule *kana*, resulting in a vowel change: *ja > ji, ti > te, fe > fu*, etc. In the examples shown below, the kana digraph is boxed:

(19) memory stick   → ***memo**rii+**suti**Qku*   → *memosu[te]* 'USB stick, flash memory'

      メモリー + ス[ティ]ック   メモス[テ]

  mass communication → ***masu**+**ko**myunikeešoN* → *masuko[mi]* 'mass media'

      マス + コ[ミュ]ニケーション   マスコ[ミ]

  differential gear   → ***defa**renšaru+**gia***   → *de[fu]gia* 'differential gear'

      デ[ファ]レンシャル + ギア   デ[フ]ギア

Mora-splitting is infrequent, occurring only 13 times in El Cid and accounting for 3% (13/432) of the database. Of these, 10 are 2μ+2μ types where all syllables are light. In a nonce compound survey, Irwin (2009b) concluded that mora-splitting was highly marked. The El Cid data, however, is inconclusive. If one confines output to 2μ+2μ types only, then there are 17 cases where the Japanese input contains a mora digraph as the second mora in either the first or second element, i.e. where the Japanese input is of the form $\lambda\lambda_\lambda...+\lambda\lambda...$ or $\lambda\lambda...+\lambda\lambda_\lambda...$,[26] λ indicating a reduced *kana* graph and λ_λ a *kana* digraph. Of these, 10 (59%) exhibit mora-splitting and 7 (41%) do not.

PERNICIOUS HOMOPHONY: Compared to donor English, Japanese is phonemically impoverished.[27] Consequently, during phonological adaptation (Section 1.3), several English donor phones may converge on a single Japanese phoneme: e.g. Eng. *r* and *l* both → Jp. *r*; Eng. *ʤ*, *ð* (before certain vowels) and *z* (before certain vowels) all → Jp. *j*; a range of English vowels → Jp. *a* (for greater detail see Irwin 2011a: 95–96). Homophonous English loanword elements exist prior to compound clipping – the reduction occasioned by the clipping process itself only exacerbates homophony to the extent that it may become pernicious. The case of the clipped element *koN* is particularly well-known, deriving as it does from at least 11 different unclipped English elements (Irwin 2011a: 145). Other examples include *iN* and *paN*:

---

26 There is no input of the form $\lambda\lambda_\lambda...+\lambda\lambda_\lambda...$, i.e. where a mora digraph appears as the second mora in both the first and second element.

27 Japanese has 26–28 phonemes (see Figure 1), as against 40–45 for English depending on variety and analysis.

(20) inter-college   → ***iɴ**taa+**kare**ǫji*   → *iɴkare*
      inferiority complex   → ***iɴ**ferioritii+**koɴ**pureqkusu* → *iɴkoɴ*
      instrument panel   → ***iɴ**sutorumeɴto+**pane**ru* → *iɴpane*
      natural + punch   → ***načua**ru+**paɴ**či*   → *načupaɴ* (band)
      running pants   → ***raɴ**niɴgu+**paɴ**cu*   → *raɴpaɴ*
      Smashing Pumpkins → ***suma**ǫšiɴgu+**paɴ**pukiɴsu* → *sumapaɴ* (band)

Such pernicious homophony is a source of confusion not only for foreign learners of Japanese, but also for Japanese native speakers. It does not appear to be the case that the avoidance of pernicious homophony plays a role in deviations from the unmarked 2μ+2μ output found in compound clippings.

Compound clipping is not confined to the loanword stratum, but occurs freely in the Sino-Japanese stratum also: e.g. ***too**kyoo+**dai**gaku* → *toodai* 'Tokyo University'. Although it may be assumed the process was active – or at the very least activatable – from the mid-19th century when English loans began to be borrowed (Section 1.2), in fact the earliest attestations for any of the El Cid compound clippings date as far back as only the late 1920s or early 1930s: e.g. *zenesuto* 'general strike', *omuraisu* 'rice-stuffed omelette', *moga* 'flapper', *haɴsuto* 'hunger strike' (Arakawa (ed.) 1977; NKD 2000–2002).

Finally, compound clippings have even been borrowed back into English: 'cosplay' ← *kosupure* in Table 4 being one well-known example.

3.2 Morpho-orthographic truncation

Japanese possesses three 'native' scripts. *Kanji* – Chinese characters borrowed from around the 5th century – and two moraic scripts that evolved from these characters in the 9th century, *hiragana* and *katakana*.[28] The conventional domains of use of these scripts are highly complex: see Irwin (2011a: 166–168) for a summary. In brief, though, one of the major domains of the *katakana* script is to write loanwords, and loanwords are written predominantly in *katakana* script.

Recently, the Roman alphabet has been making inroads as a loanword script, although not, as one might expect, as an alternative orthography for a full, or even a truncated, loan. Thus, <needs> for (6h) or <idling> for (6j) are not making any inroads on the conventional *katakana* orthographies of <ニーズ> and <アイドリング>. Instead, a compound containing two or more English loan elements may be abbreviated through having its orthography transposed to the original Roman script, and all but the initial letters of each element deleted. To my knowledge, no detailed description or thoroughgoing analysis of this phenomenon has been published to date. What follows is a first attempt at a synthesis.

---

28 Thus, none of these three scripts is ultimately 'native'. See Seeley (2000) for a history of Japanese writing systems.

This process – which is not obligatory – is illustrated in Figure 4, where the two English loan elements forming the input are *ofisu* 'office' and *redii* 'lady' (STAGE ❶). These are assembled (see Section 1.4) into *ofisu+redii* (STAGE ❷), the script transposed to the original English (STAGE ❸), the non-initial letters of each element deleted and the initials retained (STAGE ❹). When these initials are converted to majuscule the process yields the morpho-orthographic truncation *OL* 'female office worker' (STAGE ❺). In this example, the result is pronounced *oo+eru* (← O+L).

Figure 4: The morpho-orthographic truncation process

Two important caveats must be applied to the morpho-orthographic truncation process illustrated in Figure 4. The first is that the STAGE ❺ output does not exist in English, at least not as an acronym derived from an identical input.[29] Where it does exist and the input is identical, I assume it is a direct borrowing and the process in Figure 4 has not taken place. This is the case with, for example, *enu+jii+oo* 'NGO', *pii+dii+efu* 'PDF', *ee+ii+dii* 'AED = automated external defibrillator', *pii+tii+ee* 'PTA', *jii+dii+pii* 'GDP', *ii+tii+šii* 'ETC = electronic toll collection' or *oo+bii* 'OB', all of which exist in English and must be borrowings. Thus, although on the surface *OB* and *OL* appear orthographically similar, their underlying morphology is different. Furthermore, when the corpus of Roman alphabet acronyms found in Japanese is viewed as a whole (Irwin 2011a: 188–191), the former *OB*-type direct borrowings significantly outweigh the latter *OL*-type morpho-orthographic truncations.

The second caveat is that the process in question is not one of translation. For example, the names of political parties in Japan all possess Roman letter acronyms: LDP, DPJ, JCP, etc. It is presumed these were created in Japan and are direct translations of the their Japanese originals (*jimiNtoo* 自民党 'Liberal Democratic Party', *miNšutoo* 民主党 'Democratic Party of Japan', *nihoN kyoosaNtoo* 日本共産党 'Japanese Communist Party', etc.). These acronyms are not, however, used in Japanese and cases such as these will not be considered as instances of morpho-orthographic truncation. Grey areas nevertheless exist, notably with the names of corporations. For example, *JNR* (← <japan.national.railways>), the precursor of *JR* (← <japan.railways>), dates back to 1948, but is likely – though I have no proof – the American GHQ translation of the Japanese name *niQpoN kokuyuu tecudoo* 日本

---

29 i.e. although OL does indeed exist in English, it does so as an acronym for 'offensive line' in American football, not as an acronym for 'office lady'.

国有鉄道. The multinational NEC Corp. has a long history dating back to 1899, but whether *NEC* is a translation of the Japanese corporate name *niꞯpoN deNki* 日本電気 or an original morpho-orthographic truncation (← <nippon.electric.corporation>) is unclear. Such problematic acronyms will, *faute de mieux*, be considered as morpho-orthographic compounds in the discussion to follow.

Several other issues concerning the process schematized in Figure 4 require further elucidation. These are formalized in Table 5, along with examples:

OMISSION OF STAGE ❶ Morpho-orthographic truncation may also be applied to imported compounds. In such cases, STAGE ❶ is omitted: e.g. <home.page> → *HP* or <commercial.message> → *CM*.

HYBRID INPUT STAGE ❶ may contain numerals: e.g. <basement.2> (see Table 5), <3.livingroom.diningroom.kitchen> → *3LDK* '(apartment with) three bedrooms, a living room and a dining room/kitchen'. STAGE ❶ may also contain a non-loan element: e.g. <nippon.telegraph.telephone> → *NTT*, <all.nippon.airways> → *ANA*. The element <nippon> 'Japan' is commonly found in company names.

INCOMPLETE PROCESSING Here, one of the elements in the STAGE ❶ input fails to proceed to STAGE ❷ and remains written, unabbreviated, in native script: e.g. *jee+**araato*** ← Japan+**alert'** (written <Jアラート>) 'National Disaster Warning System'.[30] Incomplete processing may be combined with **HYBRID INPUT** (see above): e.g. *daburyuu+hai* (written <W杯>) '(soccer) World Cup', where the English loan input is the initial element *waarudo* (← world) and the incompletely processed input the final element Sino-Japanese *hai* 'cup' written in *kanji*.

PRONUNCIATION STAGE ❺ output may be pronounced in one of two ways: (i) as an alphabet string, where the letters of the Roman alphabet are read according to their phonologically adapted English names (e.g. *ee-bii-šii-dii-ii* for A-B-C-D-E, etc.);[31] or (ii) as if they were an English word. Option (i) is the unmarked pronunciation: e.g. *oo+eru* for OL or *tii+bii+esu* for *TBS*.[32] Examples of option (ii) include *mekusuto* for *MEXT* (see Table 5) and *mosu* for *MOS(Burger)* (← <mountain.ocean.sun>). With compounds containing numerals, the numeral may be read as either the English loan (***waN**+eru+dii+kee* for *1LDK* ← <1.livingroom.diner.kitchen>, '(apartment with) one bedroom, a living room and a diner-kitchen') or as Japanese (*bii+**iči*** for *B1* ← <basement.1>, 'lower ground floor').

---

30 One could also argue here that *japaN* (← Japan) is employed so frequently it has become regular ized as the quasi prefix *J*, read *jee*, meaning 'Japan(ese), national, nationwide'.

31 Some speakers use Roman letter names adapted from other donor languages, most especially German: e.g. *dee* for D.

32 The phonologically adapted English values of Roman letters are overwhelmingly bimoraic, or one foot. See the discussion in Section 3.1.

SPELLING   At STAGE ❸, the script is always transposed to the original English spelling, never one of the standard romanization systems employed for writing Japanese (e.g. Hepburn or *kunrei*). Thus, in Figure 4, the second element レディー is not transposed to *redii* (or the like),[33] leading to a STAGE ❹ <o+r>, and thus yielding the incorrect STAGE ❺ output *OR, to be pronounced *oo+aaru (← O+R).[34] Knowledge of the original English spelling is therefore a prerequisite for morpho-orthographic truncation. While incorrect STAGE ❺ output produced by incomplete or flawed knowledge of original English spelling cannot be ruled out, I am unaware of any examples. Although found only infrequently, some morpho-orthographic truncations may also be written in native *katakana* script: e.g. スイカ for *SUICA* (see Table 5). Here, the operation of a further putative STAGE ❻ must be presumed, where the pronunciation of the STAGE ❺ alphabet string is transposed back into native script.

NON-INITIALS AND REBUSES   At STAGE ❹, non-initial letters of an element may be retained or an initial letter deleted. Such compounds are relatively uncommon: e.g. <**s**uper.**u**rban.**i**ntelligent.**ca**rd> → *SUICA* (see Table 5).[35] Further, two initials may be combined rebus-fashion. Once again, such compounds are rare: e.g. *MEXT*, where <C.S.S> → <X> (see Table 5).

There exist compounds whose derivation is opaque. For example, I have neither heard nor read LCC or PK in English: they do not exist in my variety. Yet they do (or did) exist in certain other varieties, or in technical jargon. It is unclear whether the Japanese *CG, LCC*, and *PK* are direct borrowings from CG, LCC, and PK in certain varieties and jargons and thus have not undergone morpho-orthographic truncation, or whether <computer.graphics>, <low.cost.carrier> and <penalty.kick> have independently undergone the process beginning at STAGE ❷ (see OMISSION OF STAGE ❶ above).

Morpho-orthographic truncation is a productive process and any English loan may serve as STAGE ❶ or STAGE ❷ input. Most truncations currently in use are of recent coinage, with few created before the 1970s. Ascertaining their first appearance is fraught with philological difficulties, however. Most are listed only in specialist dictionaries, such as those used for the El Cid database (Section 3.1), and few appear in corpora. The most comprehensive Japanese dictionary, NKD (2000–02), lists only a handful, and offers attestation dates for even fewer: the first written attestation for *BG* is 1960, for *CM* 1963 and for *OL* 1974. Older, now generally obsolete attestations do exist however. Umegaki (1944: 321), for example, cites *CM* (← <communist.manifesto>), *os* (← <old.style>) and *bc* (← <birth.control>), suggesting morpho-orthographic truncation

---

33 Japanese has no lateral phonemes and all donor laterals are adapted to Japanese /r/ (Irwin 2011a: 91). This /r/ is transcribed as <r> in all romanization systems.

34 In the majority of cases, however, transposition to the original English spelling and transposition to a romanization system will in fact yield the same STAGE ❺ output.

35 In the case of *SUICA*, and indeed some other railway smart cards, the output is also an intentional pun: see Hibiya (2012) for details.

Table 5: Examples of morpho-orthographic truncation

Process	STAGE ❶	STAGE ❷	STAGE ❸	STAGE ❹	STAGE ❺	pronunciation: (1) as alphabet string	pronunciation: (2) as English word	English gloss
standard process	ビジネスガール	ビジネス+ガール	business.girl	b.g	BG	bii+jii		businesswoman
	ゴールデンウイーク	ゴルデン+ウイーク	golden.week	g.w	GW	jii+daburyuu		cluster of national holidays in early May
	セックスフレンド	セックス+フレンド	sex.friend	s.f	SF	esu+efu		casual sex partner
omission of stage ❶		コミュニケーション+サテライト	communication.satellite	c.s	CS	šii+esu		satellite broadcasting system
		インターチェンジ	inter.change	i.c.	C	ai+šii		expressway junction
		パーキング+エリア	parking.area	p.a.	PA	pii+ee		expressway parking area
hybrid input	東京ブロードキャスティングシステム	どうきょう+ブロードキャスティング+システム	tokyo.broadcasting.system	t.b.s	TBS	tii+bii+esu		Tokyo Broadcasting System (TV channel)
	オール日本エアウェーズ	オール+にっぽん+エアウェーズ	all.nippon.airways	a.n.a	ANA	ee+enu+ee	ana	ANA Corp.
	ベースメント2	ベースメント+2	basement.2	b.2	B2	bii+ni / bii+cuu		2nd lower ground floor
incomplete processing	ジャパンリーグ	ジャパン+リーグ	japan.リーグ	j.リーグ	Jリーグ	jee+riigu		Japan National Soccer League
	ビクトリーゴール	ビクトリー+ゴール	victory.ゴール	v.ゴール	Vゴール	bui+gooru		golden goal (in soccer)
	ユーズドカー	ユーズド+カー	used.カー	u.カー	Uカー	yuu+kaa		second-hand car

Table 5: (continued)

	STAGE ❶	STAGE ❷	STAGE ❸	STAGE ❹	STAGE ❺	pronunciation: (1) as alphabet string	pronunciation: (2) as English word	English gloss
non-initials and rebuses	スーパーアーバンインテリジェントカード	スーパー＋アーバン＋インテリジェント＋カード	super.urban. intelligent.**card**	s.u.i.ca	SU CA		*suika*	rechargeable railway smart card
	マルティーアミニティーエクスプレス	マルティー＋アミニティー＋エクスプレス	multi.amenity. express	m.a.x	MAX		*maQkusu*	double-decker bullet train carriage
	ミニストリエデュケーションカルチャースポーツサイエンステクノロジー	ミニストリ＋エデュケーション＋カルチャー＋スポーツ＋サイエンス＋テクノロジー	ministry. education. culture. sports. science. technology	m.e.c.s.s.t	MECSST > MEXT		*mekusuto*	Japanese Ministry of Education, Culture, Sports, Science & Technology
direct borrowing (process does not apply)					NPO	*enu+pii+oo*		non-profit organization
					A DS		*eezu*	acquired immuno-deficiency syndrome
					UNESCO		*yunesuko*	United Nations Educational, Scientific & Cultural Organization

may be traced to as far back as before World War II. Note that *os* and *bc* are cited in miniscule, rather than majuscule. One may thus speculate that in early cases of morpho-orthographic truncation STAGE ❺ did not apply.

Frequency data is also problematic. Nakayama, Kiryū and Yamaguchi (2007: 377), an analysis of all Tokyo editions of the *Mainichi Shinbun* published 1994–2003, list only 68 Roman acronyms, the vast majority of which are direct borrowings and thus irrelevant. The most frequent morpho-orthographic truncations are, in order, *JR*, *NTT*, *TBS*, *NEC*, *CM*, *JT*, *BS*, *CS*, *JCO*, *IC*, *JA*, *OL*, *JAS*, *CG*, *JTB*, *JAL*, and *JICA*.[36] Of these, all but four (*CM*, *IC*, *OL*, *CG*) are corporations, organizations, or institutions, of which more than half contain an initial *J* for 'Japan'.

Of interest sociolinguistically is that taboo constraints do not apply (Allan and Burridge 2006): in any society where English is the official or majority language, taboo and quasi-taboo acronyms are scrupulously avoided.[37] This is not the case with morpho-orthographic truncation, as the examples in (21) attest. English native-speakers involved in or dealing with these entities may feel extremely uncomfortable or, when it comes to February 14th, enjoy the irony.

(21) COC   Centre of Community (Japanese Ministry of Education project)
      PIS   Public Information Style (Corporation)
      TOSS Teacher's Organization of Skill Sharing
      NIG   National Institute of Genetics
      VD   Valentine's Day

Processes broadly similar to morpho-orthographic truncation occur in other Japanese vocabulary strata. Here, the process is identical to that outlined in Figure 4, with the caveat that the STAGE ❶ input is non-loan. Examples include 女子 + 高生 → *joši.koosei* → j.k → *JK* 'highschool girl' and 空気 + 読めない → *kuuki.yomenai* → k.y → *KY* 'out of touch with reality' (lit. 'can't read the air'). Indeed, this non-loan morpho-orthographic truncation is colloquially known as *keewaišiki nihoɴgo* 'KY-style Japanese', after the second example cited. KY-style Japanese also employs rebuses, such as *majide.moo.muri* → m.m.m → *MMM > 3M* 'totally out of the question', and dates back to at least the 1960s. See Kitahara (ed.) (2008) or Blockbuster and GRK (eds.) (2008) for further examples.

Areally, morpho-orthographic truncation is not limited to Japanese: elsewhere in East Asia, the phenomenon occurs in Chinese (Liu 2001, 2002) and Korean (Lee and

---

36 No definitions are given by Nakayama et al. While it is safe to presume that, for example, *JTB* refers only to JTB Corp. (← <japan.travel.bureau>), this is not necessarily the case with others, such as *IC* (see Table 5), which may also refer to direct borrowings with an identical structure (← IC 'integrated circuit'). Their frequency ranking should thus be treated with caution.

37 Unless as a deliberate marketing ploy: e.g. the quasi taboo branding FCUK for 'French Connection UK'.

Ramsey 2000: 3; Kim-Renaud 2012: 164–165) and quite probably, though as yet undo-cumented, in other languages employing non-Roman script.

3.3 Ellipsis

The third and final type of truncation found in English loanword compounds is ellipsis. Here, an entire element, either the final (22a) or the initial (22b), is deleted. The former process is considerably more common than the latter. There are also a few rare cases where the middle element in a three-element English loanword is de-leted: this is illustrated in (22c). The loan input may be either imported or assembled.

(22) a. superimpose → ***suupaa****+iNpoozu* → *suupaa* 'subtitles'
         magic + pen → ***majiQku****+peN* → *majiQku* 'marker'
         half-caste → ***haafu****+kaasuto* → *haafu* 'person of mixed race'

      b. screwdriver → *sukuryuu+****doraibaa*** → *doraibaa* 'screwdriver (tool)'
         game + cube → *geemu+****kyuubu*** → *kyuubu* '(Nintendo) GameCube'
         sewing machine → *sooingu+****mišiN*** → *mišiN* 'sewing machine'

      c. must-have + item → ***masuto****+habu+****aitemu*** → *masutoaitemu* 'must-have'
         soft + ice cream → ***sofuto****+aisu+****kuriimu*** → *sofutokuriimu* 'ice cream in a cone'
         ballpoint pen → ***booru****+poiNto+****peN*** → *boorupeN* 'pen'

Differentiating ellipsis from mora-clipping (14) may be problematic in cases where an imported English loan compound is opaque; i.e. where its status as a compound is unclear to a Japanese speaker. This is often the case where the loan in question is actually a prefix+base (rather than a compound) and the prefix is infrequent:

(23) software → ***sofuto****(+)wea* → *sofuto*
     infrastructure → ***iNfura****(+)sutorakučaa* → *iNfura*
     forehand → ***foa****(+)haNdo* → *foa*

Finally, as with compound clipping (20), ellipsis can give rise to pernicious homo-phony:

(24) superimpose → ***suupaa****+iNpoozu* → *suupaa* 'subtitles'
     supermarket → ***suupaa****+maakeQto* → *suupaa* 'supermarket'
     superheterodyne → ***suupaa****+heterodaiN* → *suupaa* 'superheterodyne receiver'
     software → ***sofuto****(+)wea* → *sofuto* 'software'
     soft + ice cream → ***sofuto****+aisu+kuriimu* → *sofuto* 'ice cream in a cone'
     softball → ***sofuto****+booru* → *sofuto* 'softball'

4 Conclusions and future research perspectives

This chapter has presented a comprehensive survey of the morphophonological and semantic properties of English loanwords in terms of the word formation processes in which they participate. It is hoped that future work on the morphology of English loanwords will focus on the various truncation processes described in Section 3. More specifically:

– The high fluidity witnessed in the truncation patterns illustrated in Table 3 requires thorough diachronic and synchronic analysis. What motivates the skewing towards a particular truncation type? Are the factors involved sociolinguistic, phonological, morphophonological, suprasegmental, a combination of several of these, or something else entirely?

– The operation of these same processes appears to require an awareness by the Japanese speaker that the object of truncation is a compound, despite in many cases the lack of any linguistic signal to this effect. The operation of morpho-orthographic truncation implies a further awareness of English spelling. These observations throw up a number of interesting issues in need of further research.

– Possible motivations proposed in Section 3.1 for the prevalence of 2μ+1μ type clipped compounds (Table 4, ex. (15)) are inadequate. The same applies to those offered for NON-FINAL VOWEL DELETION IN NON-LIGHT SYLLABLES and for DELETION OF THE MORA OBSTRUENT Q. Further research is required to bring to light these unknown, probably sociolinguistic, variables.

– Morpho-orthographic truncation is a new, under-researched, highly fluid phenomenon. The description offered in Section 3.2 is tentative and painted with a broad brush. More data collection is required and any future corpus-based analyses are certain to throw up theoretically important hypotheses and results.

Acknowledgments

References

Allan, Keith and Kate Burridge. 2006. *Forbidden words: Taboo and the censoring of language.* Cambridge: Cambridge University Press.

Arakawa, Sōbei (ed.). 1977. *Gairaigo jiten* [Loanword dictionary]. Tokyo: Kadokawa Shoten.

Arisaka, Hideo. 1940. *On'inron* [Phonology]. Tokyo: Sanseidō.

Atkinson, Hoffman. 1879. *Revised and enlarged edition of exercises in the Yokohama dialect*. Yokohama: Japan Gazette Office.

Blockbuster and GRK = Gendai Ryakugo Kenkyūkai [Research Group on Modern Abbreviations] (eds.). 2008. *KYgo jiten* [Dictionary of KY words]. Tokyo: Byakuya Shobō.

Daniels, F. J. 1948. The vocabulary of the Japanese ports lingo. *Bulletin of the School of Oriental and African Studies* 12. 805–823.

Daulton, Frank. 2008. *Japan's built in lexicon of English based loanwords*. Clevedon: Multilingual Matters.

Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.

Fukuzawa, Haruka and Mafuyu Kitahara. 2005. Ranking paradoxes in consonant voicing in Japanese. In Jeroen van de Weijer, Kensuke Nanjō and Tetsuo Nishihara (eds.), *Voicing in Japanese*, 105–121. Berlin: Mouton de Gruyter.

GJH = Gakken Jiten Henshūbu [Gakken Dictionary Editorial Board] (eds.). 2001. *Pāsonaru eiwa katakanago jiten* [Personal English-Japanese and *katakana* word dictionary]. Tokyo: Gakushūkenkyūsha.

Hamano, Shoko. 1998. *The sound symbolic system of Japanese*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Hargraves, Orin. 2003. *Mighty fine words and smashing expressions*. Oxford: Oxford University Press.

Hashimoto, Waka. 2006. *Asahi Shinbun shasetsu no gairaigo* [Loanwords in *Asahi Shinbun* editorials]. *Dōshisha Kokubungaku* [Doshisha University Journal of Japanese Literatures] 64. 178–186. Dōshisha University.

Hibiya, Junko. 1995. The velar nasal in Tokyo Japanese: A case of diffusion from above. *Language Variation and Change* 7. 139–152.

Hibiya, Junko. 1996. Denasalization of the velar nasal in Tokyo Japanese: Observations in real time. In Gregory Guy, Crawford Feagin, Deborah Schiffrin and John Baugh (eds.), *Towards a social science of language. Papers in honor of William Lavov, volume I: Variation and change in language and society*, 161–171. Amsterdam: John Benjamins.

Hibiya, Junko. 1999. Variationist sociolinguistics. In Natsuko Tsujimura (ed.), *The Handbook of Japanese linguistics*, 101–120. Oxford: Blackwell.

Hibiya, Junko. 2012. Kotoba no tayōsei ni me o hiraku [Looking at the diversity of words]. In Junko Hibiya (ed.), *Hajimete manabu shakaigengogaku* [Introduction to sociolinguistics], 1–13. Kyoto: Mineruva Shobō.

Hirai, Mitsuko. 2003. Waseieigo no keitairon [The morphology of made-in-Japan English]. *Journal of Rehabilitation and Health Sciences* 1. 54–59.

Hirata, Yukari. 1990. Perception of geminated stops in Japanese word and sentence levels. *Bulletin of the Phonetic Society of Japan* 194. 23–28.

Hirozane, Yoshito. 1992. Perception by Japanese speakers of some English sounds as the Japanese choked sound /Q/. *Bulletin of the Phonetic Society of Japan* 201. 15–19.

Ichikawa, Sanki. 1929. *Foreign influences in the Japanese language*. Tokyo: Japanese Council, Institute of Pacific Relations.

Ichikawa, Sanki. 1930. The pronunciation of English loan-words in Japanese. In N. Bøgholm, Aage Brusendorff and C.A. Bodelsen (eds.), *A grammatical miscellany offered to Otto Jespersen on his seventieth birthday*, 179–190. Copenhagen: Levin and Munksgaard.

Inoue, Aya. 2007. Grammatical features of Yokohama pidgin Japanese: Common characteristics of restricted pidgins. In Naomi Hanaoka McGloin and Junko Mori (eds.), *Japanese/Korean Linguistics* 15. 55–66. Stanford: CSLI Publications.

Irwin, Mark. 2009a. Prosodic size and rendaku immunity. *Journal of East Asian Linguistics* 18. 179–196.

Irwin, Mark. 2009b. Mora splitting in loanword compounds. *Yamagata Daigaku Jinbungakubu Kenkyū Nenpō* [Yamagata University Faculty of Literature and Social Sciences Annual Research Report] 6. 71–85. Yamagata University.

Irwin, Mark. 2011a. *Loanwords in Japanese*. Amsterdam: John Benjamins.

Irwin, Mark. 2011b. Mora clipping of loanwords in Japanese. *Journal of Japanese Linguistics* 27. 71–81.

Ishiwata, Toshio. 1976. *Waseieigo to kokusai tsūyōgo* [Made-in-Japan English and current international language]. In Bunkachō (ed.), *Gairaigo* [Loanwords], 65–73. Tokyo: Bunkachō Bunkabu Kokugoka.

Itô, Junko. 1990. Prosodic minimality in Japanese. In Michael Ziolkowski, Manuela Noske and Karen Deaton (eds.), *Papers from the 26th Regional Meeting of the Chicago Linguistic Society Volume 2: The Parasession on the Syllable in Phonetics and Phonology*, 213–239. Chicago: Chicago Linguistic Society.

Itô, Junko and Armin Mester. 1992. *Weak layering and word binarity*. Report LRC-92-09, University of California, Santa Cruz. Republished (2003) in Takeru Honma, Masao Okazaki, Toshiyuki Tabata and Shin-ichi Tanaka (eds.), *A new century of phonology and phonological theory: A Festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*, 26–65. Tokyo: Kaitakusha.

Itô, Junko and Armin Mester. 1995. Japanese phonology. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 817–838. Cambridge, MA: Blackwell.

Itô, Junko and Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell.

Itô, Junko and Armin Mester. 2015. Word formation and phonological processes. In Haruo Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 363–395. Berlin: De Gruyter Mouton.

Itô, Junko, Yoshihisa Kitagawa and Armin Mester. 1996. Prosodic faithfulness and correspondence: Evidence from a Japanese argot. *Journal of East Asian Linguistics* 5. 217–294.

Itō, Masamitsu. 2003. Kayōkyoku no naka no gairaigo gaikokugo [Loanwords and foreign words in pop songs]. *Nihongogaku* 22(7). 40–48.

JKS = Jiyūkokuminsha [Free Citizens Company] (ed.). 1977. *Gendai yōgo no kiso chishiki 1977* [Essential modern vocabulary 1997]. Tokyo: Jiyūkokuminsha.

JKS = Jiyūkokuminsha [Free Citizens Company] (ed.). 2003. *Gendai yōgo no kiso chishiki 2003* [Essential modern vocabulary 2003]. Tokyo: Jiyūkokuminsha.

JKS = Jiyūkokuminsha [Free Citizens Company] (ed.). 2010. *Gendai yōgo no kiso chishiki 2010* [Essential modern vocabulary 2010]. Tokyo: Jiyūkokuminsha.

JKS = Jiyūkokuminsha [Free Citizens Company] (ed.). 2013. *Gendai yōgo no kiso chishiki 2013* [Essential modern vocabulary 2013]. Tokyo: Jiyūkokuminsha.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Vocabulary strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Katayama, Motoko. 1995. Loanword accent and minimal reranking in Japanese. *Phonology at Santa Cruz* 4. 12–28. University of California, Santa Cruz.

Kawahara, Shigeto. 2011a. Aspects of Japanese loanword devoicing. *Journal of East Asian Linguistics* 20. 169–194.

Kawahara, Shigeto. 2011b. Japanese loanword devoicing revisited: A rating study. *Natural Language & Linguistic Theory* 29. 705–723.

Kawahara, Shigeto. 2012. Lyman's Law is active in loanwords and nonce words: Evidence from naturalness judgment studies. *Lingua* 122. 1193–1206.

Kim-Renaud, Young-Key. 1997. *The Korean alphabet*. Honolulu: University of Hawai'i Press.

Kindaichi, Haruhiko. [1942] 1967. Gagyō bionron [On the velar nasal]. In Kindaichi Haruhiko (ed.), *Nihongo on'in no kenkyū* [Studies in Japanese phonology], 168–197. Tokyo: Tōkyōdō.

Kitahara, Yasuo (ed.). 2008. *KY-shiki nihongo* [KY-style Japanese]. Tokyo: Taishukan Shoten.

Koo, John and Yayoi Homma. 1989. Consonant gemination in English loanwords of Japanese. *Journal of the Association of Teachers of Japanese* 23. 125–134.

Kubozono, Haruo. 1995. *Gokeisei to on'in kōzō* [Word formation and phonological structure]. Tokyo: Kurosio Publishers.

Kubozono, Haruo. 1999a. *Nihongo no onsei* [The sounds of Japanese]. Tokyo: Iwanami Shoten.

Kubozono, Haruo. 1999b. Mora and syllable. In Natsuko Tsujimura (ed.), *The Handbook of Japanese linguistics*, 31–61. Oxford: Blackwell.

Kubozono, Haruo. 2001a. On the markedness of diphthongs. *Kobe Papers in Linguistics* 3. 60–73. Kobe University.

Kubozono, Haruo. 2001b. Epenthetic vowels and accent in Japanese: Facts and paradoxes. In Jeroen van de Weijer and Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*, 113–142. Berlin: Mouton de Gruyter.

Kubozono, Haruo. 2002. Prosodic structure of loanwords in Japanese: Syllable, structure, accent and morphology. *Journal of the Phonetic Society of Japan* 6(1). 79–97.

Kubozono, Haruo. 2015. Loanword phonology. In Haruo Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 313–361. Berlin: De Gruyter Mouton.

Kuwamoto, Yūji. 1998a. Ryakugokeisei kara mita nihongo no chōjūonsetsu no kōzō ni tsuite [Truncation and the structure of trimoraic syllables in Japanese]. *Gengo Kagaku Ronshū* [Papers in Linguistic Science] 2. 25–35. Tohoku University.

Kuwamoto, Yūji. 1998b. Nihongo ni okeru fukugōgoryakugo no on'inkōzō [The phonological structure of Japanese clipped compounds). *On'in Kenkyū* 1. 161–168.

Labrune, Laurence. 2002. The prosodic structure of simple abbreviated loanwords in Japanese: A constraint-based account. *Journal of the Phonetic Society of Japan* 6(1). 98–120.

Labrune, Laurence. 2012a. *The phonology of Japanese*. Oxford: Oxford University Press.

Labrune, Laurence. 2012b. Questioning the universality of the syllable: Evidence from Japanese. *Phonology* 29. 113–152.

Lee, Iksop and S. Robert Ramsey. 2000. *The Korean language*. Albany: State University of New York Press.

Liu, Yongquan. 2001. New development of Chinese lexicon. *Journal of Asian Pacific Communication* 11(2). 255–262.

Liu, Yongquan. 2002. On the problems of Chinese lettered words. *Sino Platonic Papers* 116. 1–13.

Loveday, Leo. 1996. *Language contact in Japan: A sociolinguistic history*. Oxford: Oxford University Press.

Lovins, Julie Beth. 1973. *Loanwords and the phonological structure of Japanese*. Chicago, IL: University of Chicago dissertation.

Maruyama, Takao, Tadao Kobayashi, Chiaki Yamazaki and Masanori Terauchi. 1992. *Katakanago o eigo ni suru jiten* [Dictionary of *katakana* words and their English translations]. Tokyo: Taishukan Shoten.

McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Mester, Armin. 1990. Patterns of truncation. *Linguistic Inquiry* 21. 478–485.

Nakagawa, Yoshio. 1966. Rendaku, rensei (kashō) no keifu [A genealogy of sequential voicing and sequential non-voicing (working label)]. *Kokugokokubun* 35. 302–314. Kyoto University.

Nakayama, Eriko, Rika Kiryū and Masaya Yamaguchi. 2007. Shinbun ni miru kikangairaigo [Core loanwords found in newspapers]. In Kokuritsu Kokugo Kenkyūjo [National Institute for the Japanese Language] (ed.), *Hōkoku 126: Kōkyōbaitai no gairaigo   gairaigo iikae teian o sasaeru chōsa kenkyū* [Report #126: Loanwords in the public media – basic research on 'paraphrasing loanwords'], 343–378. Tokyo: Kokuritsu Kokugo Kenkyūjo.

NINJAL = Kokuritsu Kokugo Kenkyūjo [National Language Research Institute] (ed.). 1964. *Hōkoku 25: Gendai zasshi 90 shu no yōgoyōji daisanbunsatsu III: bunseki* [Report #25: Vocabulary and Chinese characters in 90 magazines of today III: Analysis of the results]. Tokyo: Shūei Shuppan.

NINJAL = Kokuritsu Kokugo Kenkyūjo [National Institute for the Japanese Language] (ed.). 1987. *Hōkoku 89: Zasshi yōgo no hensen* [Report #89: Vocabulary trends in magazines]. Tokyo: Shūei Shuppan.

NINJAL = Kokuritsu Kokugo Kenkyūjo [National Institute for the Japanese Language] (ed.). 2006. *Gairaigo iikae tebiki* [A guide to paraphrasing loanwords]. Tokyo: Gyōsei.

NKD = Shōgakukan (ed.). 2000–2002. *Nihon kokugo daijiten* [The great dictionary of the Japanese language]. Tokyo: Shōgakukan.

Ōtsuki, Fumihiko. 1884. Gairaigogen kō [Notes on loanword etymologies]. *Gakugei Shirin* 14. 122–139, 370–384, 572–590. Tokyo University.

Poser, William. 1990. Evidence for foot structure in Japanese. *Language* 66. 78–105.

Quackenbush, Hiroko. 1977. English loanwords in Japanese: Why are they difficult for English-speaking students? *The Journal of the Association of Teachers of Japanese* 12. 149–173.

Quackenbush, Hiroko. 1989. Gairaigo no sokuonka ni tsuite [Gemination in Japanese loanwords]. In Hiroshima Daigaku Kyōikugakubu [Hiroshima University Faculty of Education] (eds.), *Ryūgakusei nihongokyōiku ni kansuru rironteki jisenteki kenkyū* [Theoretical and practical research on Japanese language education for foreign students], 1–7. Hiroshima: Hiroshima Daigaku Kyōikuga-kubu Ryūgakuseinihongokyōiku / Nihongokyōiku Gakka. Hiroshima University.

Rosen, Eric Robert. 2003. Systematic irregularity in Japanese rendaku: How the grammar mediates patterned lexical exceptions. *Canadian Journal of Linguistics / Revue Canadienne de Linguistique* 48. 1–37.

Seeley, Christopher. 2000. *A history of writing in Japan*. Honolulu: University of Hawai'i Press.

Shibasaki, Hideko, Katsuo Tamaoka and Yūki Takatori. 2007. Amerikajin wa waseieigo o donogurai rikai dekiru ka [How much can Americans understand made-in-Japan English?]. *Nihongo Kagaku* 21. 89–110.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Shinmura, Izuru (ed.). 2008. *Kōjien dai 6 han*. Tokyo: Iwanami Shoten.

SSDH = Sanseidō Henshūjo [Sanseidō Editorial Institute] (ed.). 2010. *Konsaisu katakanago jiten dai 4 ban* [Concise *katakana* word dictionary: 4th edition]. Tokyo: Sanseidō.

Sugimoto, Tsutomu. 2007. Gairaigo [Loanwords]. In Yoshifumi Hida, Yoshihide Endō, Masanobu Katō, Takeyoshi Satō, Kiyoto Hachiya and Tomiyoshi Maeda (eds.), *Nihongogaku kenkyū jiten* [The research encyclopaedia of Japanese linguistics], 408–411. Tokyo: Meiji Shoin.

Suzuki, Shunji. 2008. Waseieigo no kenkyū: sono kōzō to shisō [A study of Japanized English: Its structure and thought]. *Kokusai Tankidaigaku Kiyō* [Bulletin of Kokusai Junior College] 23. 1–47. Kokusai Junior College.

Takagi, Naoyuki and Virginia Mann. 1994. A perceptual basis for the systematic phonological corre-spondences between Japanese load (sic.) words and their English source words. *Journal of Phonetics* 22. 343–356.

Takayama, Tomoaki. 2005. A survey of rendaku in loanwords. In Jeroen van de Weijer, Kensuke Nanjo and Tetsuo Nishihara (eds.), *Voicing in Japanese*, 177–190. Berlin: Mouton de Gruyter.

Tamaoka, Katsuo, Mutsuko Ihara, Tadao Murata and Hyunjung Lim. 2009. Effects of first-element phonological length and etymological-type features on sequential voicing (*rendaku*) of second elements. *Journal of Japanese Linguistics* 25. 17–38.

Tanabe, Yōji. 1990. Waseieigo no keitaibunrui [A morphological categorization of made-in-Japan English]. *Waseda Daigaku Nihongokenkyūkyōiku Sentā Kiyō* [Bulletin of Waseda University Center for Japanese Language] 2. 1–26. Waseda University.

Tateishi, Koichi. 1989. Theoretical implications of Japanese musicians' language. In Jane Fee and Kathryn Hunt (eds.), *Proceedings of the eighth West Coast Conference on Formal Linguistics*, 384–398. Stanford: CSLI Publications.

Tōdō, Akiyasu. 1969. *Kango to nihongo* [Sino-Japanese and Japanese]. Tokyo: Shūei Shuppan.

Tranter, Nicolas and Mika, Kizu. 2012. Modern Japanese. In Nicolas Tranter (ed.), *The languages of Japan and Korea*, 268–312. Abingdon: Routledge.

Tsujimura, Natsuko. 2007. *An introduction to Japanese linguistics*. Oxford: Blackwell.

Tsujimura, Natsuko and Stuart Davis. 1987. The accent of long nominal compounding in Tokyo Japanese. *Studies in Language* 11(1). 199–206.

Ueda, Kazutoshi, Junjirō Takakusu, Kurakichi Shiratori, Naojirō Murakami and Shōzaburō Kanazawa. 1915. *Nihon gairaigo jiten* [Dictionary of Japanese loanwords]. Tokyo: Sanseidō Shoten.

Umegaki, Minoru. 1944. *Nihongairaigo no kenkyū zōhoban* [A study of Japanese loanwords: Revised and enlarged edition]. Osaka: Seinentsūshinsha.

Umegaki, Minoru. 1956. *Ingo jiten* [Dictionary of Cant, Argot, and Secret Language]. Tokyo: Tōkyōdō.

Umegaki, Minoru. 1963. *Nihongairaigo no kenkyū* [A study of Japanese loanwords]. Tokyo: Kenkyusha.

Vance, Timothy. 1987. *An introduction to Japanese phonology*. Albany: State University of New York Press.

Vance, Timothy. 2008. *The sounds of Japanese*. Cambridge: Cambridge University Press.

Vance, Timothy. 2015. Rendaku [Sequential voicing]. In Haruo Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 397–441. Berlin: Mouton de Gruyter.

Watanabe, Toshirō, Edmund Skrzypczak and Paul Snowden (eds.). 2003. *Kenkyusha's new Japanese English dictionary*. 5th edition. Tokyo: Kenkyusha.

Yamada, Yūichirō. 2005. *Gairaigo no shakaigaku* [The sociology of Japanese loanwords]. Yokohama: Shunpūsha.

Yamaguchi, Toshiko. 2007. *Japanese linguistics: An introduction*. London: Continuum.

Yonekawa, Akihiko. 1997. *Wakamono kotoba jiten* [Dictionary of youth speech]. Tokyo: Tōkyōdō.

Yuzawa, Tadayuki. 2010. *Kodai nihonjin to gaikokugo* [The ancient Japanese and foreign languages]. Tokyo: Bensei Shuppan.

Takayasu Namiki and Taro Kageyama
6 Word structure and headedness

1 Introduction

This chapter addresses the questions of how complex words in Japanese are built from two or more morphemes/words and how their morphological structure is associated with their meaning. It is a fundamental task of morphological investigations to account for how the category of a complex word is determined from the lexical information of its constituents, where the notion "head" plays a pivotal role. In syntax, the head of a phrase is the X^0 (word) element that determines the category of a whole phrase containing it, and the position of a phrasal head varies in individual constructions and in individual languages. In an adjectival modification construction like *a tall building*, the noun *building* is the head because it determines the category of the whole noun phrase, and so is the same noun in the relative clause construction like *a building that is tall*. While the mainstream approach in generative-oriented morphology holds that the identification of heads in word structure could in principle be approached by the same methodology as in syntax (cf. Selkirk 1982; Di Sciullo 2005), quite a few complications emerge in the realm of morphology from such lexical idiosyncrasies as bound roots whose categorial status is not straightforwardly determined, grammaticalized morphemes whose categories have become obscure, and unique behaviors of individual vocabulary strata.

The notion "head of a word" can be traced back at least to Bloomfield's (1933) distinction between endocentric compounds (compounds that have a categorial or semantic center inside, such as *blackbird* 'a kind of bird' and *door-knob* 'a kind of knob') and exocentric compounds (compounds that do not, such as *gadabout* 'one who gads about' and *turnkey* 'jailer'). Bloomfield's conception of the head thus appears to build on morphological (categorial) and semantic considerations. Since then, the complexity of the factors involved in word structure has given rise to a variety of approaches to the characterization of "head of a word" such as a purely categorial approach (Williams 1981), a purely semantic approach (Haspelmath 2002), a mixed approach referring to both category and meaning (Aronoff and Fudeman 2005), and a multifaceted approach resorting to category, meaning, and morphological features (Scalise, Fábregas, and Forza 2009). There is even a view that repudiates the feasibility of such a notion (Bauer 1990). In this chapter, we will begin by examining perhaps the most influential work in this area, namely Williams (1981), and then discuss the relevance of semantic factors to the determination of categories.

Williams (1981) proposed the Right-hand Head Rule (RHR) to pinpoint the head position of word structure uniformly by linear order in both compounds and affixed words: "The head of a complex word is the right-hand member of that word" (p. 248). Observe the examples in (1).

(1) a. Compounds: [[*dark*]$_A$ [*room*]$_N$]$_N$, [[*tax*]$_N$ [*free*]$_A$]$_A$

   b. Prefix: [*re* [*construct*]$_V$]$_V$

   c. Suffix: [[*construct*]$_V$ *ion*]$_N$]$_N$

In (1a), *darkroom* is a compound noun because *room* on the right is a noun, and *tax-free* is a compound adjective because *free* on the right is an adjective. In (1b), since *construct* on the right is a verb, *reconstruct* is also a verb. In (1c), the bound suffix *-ion* is treated as "N" because, according to Williams, the verb *construct* changes to a noun by virtue of this suffix, and hence *-ion* itself bears the category "N". The innovation of Williams's theory thus lies in the extension of category labels from independent words to bound affixes. The assignment of categories to bound morphemes can be extended to the bound stems of inflecting words. For example, although the stem *huru-* 'old' of the adjective *huru-i* [old-PRS] cannot be used as an independent word, we can assign the category "A" to the stem *huru-* itself because the inflectional ending *-i* is irrelevant to category determination; likewise, we can assign the category "V" to the stem *hasir-* 'run' of the verb *hasir-u* [run-PRS].

It is immediately clear that the categorial head (*room*) of the compound noun *darkroom* also serves as the semantic core that characterizes the meaning of the whole word. Using the hyponymy relation "*X-Y* is a kind of *Y*" (Allen 1978; Haspelmath 2002), we can say that *darkroom* is a kind of *room*, and hence *room* is the "semantic head" of *darkroom* as well as its categorial head. The semantic notion of hyponymy, however, is not easily applicable to the other cases in (1). It is hard to say that *tax-free* is a kind of *free* or *construction* is a kind of *-ion*. Williams' RHR has the advantage of eschewing such semantic issues by limiting the definition of head exclusively to the categorial information. Nevertheless, semantic considerations should not be belittled because strict application of the RHR would lead to the counter-intuitive conclusion that the noun *turnkey* 'jailer', an example of exocentric compound, should be analyzed as an endocentric compound whose head is the noun *key* in the right-hand position. These examples suggest that neither category nor meaning is, by itself, sufficient to characterize the notion of "head of a word".

In light of these considerations, we will basically follow the categorial definition of a head while at the same time paying due attention to semantic factors in determining category structure. Williams' rigid delimitation of a categorial head as the right-hand element has been confronted with a number of counterexamples within a single language and across different languages, and it is now generally regarded as "at best a language-specific principle" (Booij 2012: 56). In fact, Kageyama (1982, 2009) observed for Japanese that, although the right-headed structure pertains to a good majority of native complex words, other kinds of structure – left-headed, double-headed, and headless (i.e. exocentric) – are also available, notably in the Sino-Japanese stratum. This chapter will thus survey the distributions of the four

types of head structure in Japanese and bring to light interesting interactions of category determination and semantic interpretations.

Our discussion will proceed as follows. Section 2 surveys the headed (endocentric) structures in Japanese by dividing them into right-hand head, left-hand head, and double head structures, whereas Section 3 focuses on headless or exocentric structures. In Section 4, a close inspection of headedness leads us to question the validity of the distinction between compounds and derived words. Section 4 will also bring out concrete instances of morphological change from words to affixes and change from affixes to words. Through these discussions it will become apparent that semantic factors make a non-trivial contribution to the determination of word structure.

2 Headed structure

This section will discuss the headed word structures in Japanese morphology by dividing them into right-headed structure (2.1), left-headed structure (2.2), and double-headed structure (2.3).

2.1 Right-headed structure

There is no doubt that the basic word structure of Japanese is right-hand headed in line with the head-final parameter in syntax, according to which adjectival modifiers and relative clauses, for example, consistently precede their head nouns in noun phrases. In fact, the right-headed structure is observed in an overwhelming number of compound nouns coming from the native Japanese stratum, as illustrated in (2).

(2) a. *naki-gao* [cry$_V$-face] 'a tearful face'

b. *huru-gao* [old$_A$-face] 'an old face, an old member'

c. *hen-gao* [funny$_{AN}$-face] 'a funny face intentionally disfigured to make people laugh'

d. *doya-gao* [see.how.I.did.it$_{EXCL}$-face] 'a self-satisfied look'

e. *ma-gao* [serious$_{PREF}$-face]] 'a serious look, a straight face'

f. *yoko-gao* [side$_N$-face] 'a profile, a side view of a face'

In the examples of (2), the left-hand constituents diverge in categories as shown by the category names appended to them, whereas the right-hand members are consistently manifested by the noun *kao* 'face, look', phonologically realized as *gao*

due to the *rendaku* rule of consonant voicing inside compound nouns. Inasmuch as all these compounds invariably function as nouns despite the categorial diversity of the left-hand constituents, the right-hand noun *kao* 'face, look' is identified as their categorial head.

Here a query might be raised about the last example *yoko-gao* [side-face] 'profile' in (2f). Since this compound consists of two nouns, it is not immediately clear whether the category of the entire compound is determined by *yoko* 'side' or *gao* 'face'. It is in such cases where category information alone is insufficient that semantic considerations are brought into play. Semantically, *yoko-gao* 'profile' represents the face of a person as seen from the side, and this manner of semantic interpretation whereby the left-hand constituent is construed as modifying the meaning of the right-hand constituent holds true of all the other compounds in (2). Assuming that all the six complex words in (2) share the same pattern, the right-hand noun *kao* 'face' is looked upon as the categorial head in (2f) as well.

An even more difficult problem is posed by compounds consisting of bound morphemes – a situation typically found with Sino-Japanese lexical items. How is it possible to know the lexical category of the bound morphemes that compose Sino-Japanese words when they cannot occur independently in syntactic structure? For example, what enables us to say that the Sino-Japanese VN *nyuu-sya* [enter-company] 'enter a company, get a job with a company' is composed of a verbal morpheme (*nyuu* 'enter') followed by a nominal morpheme (*sya* 'company')? There are several clues to recognizing their category and meaning. One is the paradigmatic relations of a set of Sino-Japanese words that take the structure "*nyuu*-X", as in *nyuu-gaku* 'enter a school', *nyuu-too* 'enter a political party', and *nyuu-situ* 'enter a room'. Since these words all fall under the general pattern of two-morpheme S-J verbal nouns taking the construction of "Chinese verb + Chinese noun" (see Chapter 3 [Kobayashi, Yamashita, and Kageyama, this volume]), the first morpheme *nyuu* 'enter' in *nyuu-sya* can be identified as a verbal element and the second morpheme *sya* 'company' as a noun element. In fact, although *sya* 'company' cannot stand alone, it acquires wordhood when it is appended with the honorific prefix *on-* 'your honorable', as in *on-sya* 'your company'. A more realist clue is the *kanji* representation. The two Chinese characters 入社 for *nyuu-sya* immediately allow the reader who knows these characters to recognize the first morpheme *nyuu* 入 as having a verbal function meaning 'enter' and the second morpheme *sya* 社 as having a nominal function meaning 'company'. As Nagano and Shimada (2014) argue, *kanji* characters largely have the function of representing Sino-Japanese lexemes.

Now, the predominance of the right-hand head as observed in (2) extends to other categories than nouns and other lexical strata than the native Japanese. Table 1 shows typical examples of right-headed compounds classified in terms of different lexical categories and lexical strata.

Table 1: Right-hand headed compounds

	native heads	Sino-Japanese heads	foreign heads
compound nouns	N-N *hidari te* [left-hand] 'left hand', V-N *yaki niku* [broil-meat] 'grilled meat', A-N *naga banasi* [long-talk] 'long chat'	N-N *horaa eiga* [horror-movie] 'horror movie', VN-N *oodan hodoo* [cross-pavement] 'pedestrian crossing'	N-N *si basu* [city-bus] 'municipal bus', AN-N *otegaru resipi* [easy-recipe] 'easy recipe'
compound verbs	N-V *ti basiru* [blood-run] 'get bloodshot', V-V *tobi oriru* [jump-get.down] 'jump down', A-V *waka gaeru* [young-return] 'be rejuvenated'	D.N.A. (There are no S-J morphemes that exhibit verb inflections.)	D.N.A. (There are no foreign morphemes that exhibit verb inflections.)
compound VNs	AN-VN *tokubetu kasidasi* [special-lending]	AN-VN *kinkyuu tyakuriku* [emergency-landing]	N-VN *zyoonai anaunsu* [on.the.spot-announcement]
compound adjectives	N-A *kyoomi bukai* [interest-deep] 'interesting', V-A *musi atui* [steam-hot] 'muggy', A-A *zuru gasikoi* [sly-clever] 'cunning'	D.N.A. (There are no S-J morphemes that exhibit adjectival inflections.)	D.N.A. (There are no foreign morphemes that exhibit adjectival inflections.)
compound ANs	N-AN *keiken yutaka* [experience-rich] 'experienced'	VN-AN *saigen kanoo* [reproduce-possible] 'reproducible'	Apparently no example.

As seen from Table 1, the native stratum presents the richest source of head identification in all the major lexical categories: N, V, A, VN (verbal noun), and AN (adjectival noun).

Before moving on to affixes, we will touch on two phenomena that are relevant to clarification of the relation between a categorial head and its semantic impact. Consider first the recursive expansion of compound nouns in (3).

(3) a. [[*moosi-komi*]$_N$ [*uke-tuke*]$_N$]$_N$
     application   receipt
     'receipt of application'

   b. [[[*moosi-komi*]$_N$ [*uke-tuke*]$_N$] [[*simekiri*]$_N$ [*bi*]$_N$]$_N$]$_N$
      application   receipt   closing   day
      'deadline for the receipt of application'   (Nishio 1976)

   c. [[[*moosi-komi*]$_N$ [*uke-tuke*]$_N$]$_N$ [*kanoo*]$_{AN}$]$_{AN}$ *na*
      application   receipt   possible   ENDING
      'possible to receive an application'

A notable feature of such multiply-embedded compounds is that a hyponym relation always holds between a whole compound and its rightmost constituent (Namiki 2001). In (3a), it is *uke-tuke* 'receipt' on the right hand that regulates the semantic interpretation of the whole compound *moosi-komi uke-tuke* 'receipt of application' as a hyponym of *uke-tuke* 'receipt'. The same mechanism of semantic interpretation applies to a longer compound in (3b), where the rightmost constituent *simekiri-bi* [closing-day] 'deadline' designates the date beyond which the application cannot be received. Thus, although the individual constituents of the compounds in (3a) and (3b) are indistinguishable in terms of lexical category, semantic consideration suggests that the compound structure is expanded to the right, and hence the right-most constituent is the categorial head. In fact, if an AN *kanoo* 'possible' is added to the right of the compound in (3a), the result is a compound AN in (3c), which as a whole takes the ending *na* when used in prenominal position.

The second phenomenon has to do with "reversible compounds" (Scalise 1992; Namiki 1994) like those in (4), where different orders of the same two nouns give rise to entirely different denotations.

(4) a. *hati-mitu*   ≠ *mitu-bati* (*hati* 'bee' becomes *bati* due to *rendaku*.)
        bee-syrup 'honey'   syrup-bee   'honeybee'

    b. *eisei-hoosoo*   ≠ *hoosoo-eisei*
        satellite-broadcasting   broadcasting-satellite

Regardless of the order differences, it is always the right-hand member that determines the semantic interpretation of a whole compound.

The right-hand head structure of (4a) is confirmed by the selection of numeral classifiers. Japanese numeral classifies are selected by the head noun they cooccur with. Thus in (4a), *hati-mitu* 'honey' is compatible with the numeral classifier *-teki* 'drop', which selects liquid nouns, as in *suu-teki no hati-mitu* [some-CLF GEN bee-syrup] 'a few drops of honey', or with the classifier *-bin* 'bottle' indicating the quantity of honey as measured by a bottle, as in *hito-bin no hati-mitu* [one-CLF GEN bee-syrup] 'a bottle of honey'. On the other hand, *mitu-bati* 'honeybee', an animate noun, goes with the numeral classifier *-hiki*, which applies to animals and insects, as in *san-biki no mitu-bati* [three-CLF GEN syrup-bee] 'three honeybees'. Since numeral classifiers are selected by the head noun, it is safe to assume that *mitu* 'syrup' is the head of *hati-mitu* 'honey' and *hati* is the head of *mitu-bati* 'honeybee'. If the selection relation between a noun and a numeral classifier is regarded as a kind of morphological agreement, then it will also count as a property of the head in word structure (cf. Scalise, Fábregas, and Forza's (2009) notion of morphological head).

The two examples of (4b) also differ in meaning according to the position of *eisei* 'satellite' and *hoosoo* 'broadcasting'. Crucially, their lexical categories differ as their meanings differ. Thus, *hoosoo-eisei* 'broadcasting satellite' is identified as a compound noun inasmuch as *eisei* 'satellite' is a plain noun, whereas *eisei-hoosoo*

'satellite broadcasting', with the VN *hoosoo* 'broadcasting' on the right, is identified as a compound VN because it can be followed by *suru* 'do'.

We are now in a position to discuss the categorial status of affixes. As mentioned in Section 1, an innovation of Williams (1981) is that a lexical category is assigned not only to independent words but also to those affixes that have a category-changing function. On this view, the suffix *-al* in *arrival* carries the category "N", deriving a noun from the verb *arrive*, but the same suffix in *autumnal* bears the category "A", deriving an adjective from the noun *autumn*. As in English, most suffixes in Japanese have a category-determining function. Table 2 shows some suffixes that function as categorial heads:

Table 2: Suffixes with head functions

	native suffixes	Sino-Japanese suffixes
I. nouns	*akaru **sa*** [bright-SUF] 'brightness', *atataka **mi*** [warm-SUF] 'warmth'	*hituyoo **sei*** [necessary-SUF] 'necessity', *syooboo **si*** [fire.fighting-SUF] 'fire fighter'
II. verbs	*uresi **gar(u)*** [glad-SUF] 'show gladness', *sizu **mar(u)*** [quiet-SUF] 'become quiet'	D.N.A. (There are no S-J suffixes that exhibit verb inflections.)
III. adjectives	*kodomo **rasii*** [child-SUF] 'childlike', *kodomo **ppoi*** [child-SUF] 'childish'	D.N.A. (There are no S-J suffixes that exhibit adjectival inflections.)
IV. ANs	*kurusi **ge*** [painful-SUF] 'painful-looking'	*kisoku **teki*** [rule-SUF] 'regular'

In the examples of Table 2, both bases and suffixes are transparent in category and meaning. In Row I, for example, *akaru-* 'bright' is the stem of a native adjective and changes to the noun *akaru-**sa*** 'brightness' by the addition of the native suffix *-sa*. Likewise, *hituyoo* 'necessary' is a Sino-Japanese AN and changes to the noun *hituyoo-sei* 'necessity' by virtue of the S-J suffix *-sei*. Because of this, the suffixes *-sa* and *-sei* are legitimately given the category "N".

The category-determining function of suffixes is also at work in lexicalized words whose internal structures can be detected only by morphological analysis, like the *cranberry* cases discussed by Aronoff (1976). Consider the lexicalized compound ANs in (5) that share the suffix *-yaka*.

(5) a.   *hare-yaka* [cloudless-SUF] 'cheerful, bright'

   b.   *maro-yaka* [round-SUF] 'mild, mellow'

   c.   *oda-yaka* [gentle-SUF] 'calm, gentle'

   d.   *karo-yaka* [light-SUF] 'light, graceful'

   e.   *koma-yaka* [fine-SUF] 'delicate, attentive'

   f.   *kirabi-yaka* [glitter-SUF] 'gorgeous'

   g.   *sito-yaka* [gentle-SUF] 'graceful'

In (5), all the bases preceding -*yaka* except *hare*- 'cloudless (sky)' in (a) are bound morphemes and can be extracted only analytically by removing the ending -*yaka*. Nonetheless, all these words, taking the -*na* ending in prenominal position, share the category of Adjectival Noun and the meaning of 'assuming the state denoted by the base'. The category-determining function should thus be ascribed to the obsolete suffix -*yaka*, which must have been productive at an earlier stage of the language.

We have so far observed that the same generalization of right-hand head holds for two qualitatively different cases. One concerns examples like those in Tables 1 and 2 whose bases and affixes are all morphologically transparent, and the other concerns examples like -*yaka* words whose bases are morphologically opaque. The distinction in morphological transparency can be roughly correlated with the productivity of the word formation process involved. Compounding in Tables 1 and suffixation in Table 2 are largely open-ended, whereas the -*yaka* words make up a closed set. In view of such a difference, the morphologically transparent word formations will be represented as productive rules yielding systematic form-meaning pairs as shown in (6), whereas the limited regularity exhibited by a set of morphologically opaque words will be captured as a construction schema like the one in (7) (cf. Booij's (2010) construction morphology model).

(6) a. $[\,\alpha\,]_X + [\,\beta\,]_Y \quad \rightarrow \quad$ form: $[\,[\,\alpha\,]_X\,[\,\beta\,]_Y\,]_Y$ (compounding)
        meaning: $\alpha\,\beta$ is a kind of β.

   a´. $[\,\alpha\,]_X + [\,suffix\,]_Y \quad \rightarrow \quad$ form: $[\,[\,\alpha\,]_X - [\,suffix\,]_Y\,]_Y$ (suffixation)
        meaning: a particular event/state/entity of α being
                 modified by the meaning of the *suffix*

(7) form: $[\,[\,\alpha\,]_X\,[\,\text{-}yaka\,]_{AN}\,]_{AN} \leftrightarrow$ meaning: [ASSUMING THE STATE OF α]

Although the majority of suffixes are category-changing, there are a few exceptions, among which is a set of socially motivated suffixes that are indispensable in Japanese culture that designate a person's title, such as -*san* (the most general title for people, attached to the surname or given name of the speaker's male or female acquaintance), -*kun* (largely limited to a male colleague or to a small boy), -*sama* (nowadays hardly used in speech but limited to the designation of an addressee on a letter), -*tyan* (applied primarily only to small children as a hypocoristic version of -*san*), and -*si* (a formal title applied to the surname of a prominent person). Like diminutive suffixes in Italian and other languages, these title suffixes are considered to have no category of their own. In addition, inflections are excluded from the discussion on headedness in word structure because they are functional categories in syntax.

In contradistinction to suffixes, prefixes are generally devoid of category-changing functions, although many of them are selective about the lexical category

of the base they attach to. Table 3 shows some examples of non-category-changing or category-neutral prefixes in boldface.

Table 3: Prefixes without head functions

	native prefixes	Sino-Japanese prefixes	foreign prefixes
I. noun bases	**ma** *natu* [true-summer] 'mid-summer', **oo** *nami* [big-wave] 'billow', **hatu** *yuki* [first-snow] 'first snowfall in the season', **o** *kane* [PREF-money] 'money'	**dai** *saigai* [big-disaster] 'major disaster', **han** *kyoosansyugi* [anti-communism], **go** *hoobi* [HON-reward]	**noo** *suriibu* [no-sleeve] 'sleeveless', **mini** *kooza* [mini-lecture], **mai** *kaa* [my-car] 'one's own car [as opposed to public transportation]'
II. verb bases	**buk** *kowasu* [violently-break] 'destroy violently'	D.N.A. (Native verbs may select only native prefixes.)	D.N.A. (Native verbs may select only native prefixes.)
III. VN bases	No example.	**tyoku** *yunyuu* [directly-import] 'import directly'	No example.
IV. adjective bases	**ko** *dakai* [slightly-high] 'slightly high', **te** *nurui* [hand-lukewarm] 'lenient'	**tyoo** *kitanai* [extra-dirty] 'very dirty'	No example.
V. AN bases	**ko** *namaiki* [little-impudent] 'cheeky'	**tyoo** *genki* [extra-healthy] 'very healthy'	No example.

The first example in Row I, *ma-* 'true', for instance, is category-neutral in that it lacks the ability to create a derived word of a new category but is merely attached to nouns (*ma-sikaku* 'true square'), adjectives (*ma-atarasii* 'brand new'), and ANs (*ma-syooziki* 'totally honest'), though not to verbs or VNs,

The category-neutral character of the prefixes like those in Table 3 will be formulated in terms of a prefixation rule like (8).

(8) $[\ prefix\ ] + [\ \beta\ ]_Y \rightarrow$ form: $[\ [prefix] - [\ \beta\ ]_Y\ ]_Y$ (prefixation)
meaning: a particular event/state/entity of β being
modified by the meaning of the *prefix*

To recapitulate, the data presented in Tables 1 to 3 confirm that the basic word structure of Japanese is right-hand headed. The prevalence of the right-headed structure in morphology is strongly correlated with the head-final structure of syntax. Prefixes, except for a handful of Sino-Japanese prefixes to be discussed shortly, do not alter the categories of the bases to which they are attached, whereas suffixes generally have a category-changing function and serve as the heads of derived

words. In compounding, the category of a whole compound is normally determined by the right-hand member, with the exception of the two-morpheme Sino-Japanese VNs discussed in Section 2.2.

2.2 Left-headed structure

In Japanese, there are two kinds of left-headed structures, both of which are arguably attributed to the patterns of Chinese morphology. One, briefly discussed in Chapter 3 (Kobayashi, Yamashita, and Kageyama, this volume), concerns Sino-Japanese compound VNs composed of a verbal element and a nominal element. In (9), three classes are distinguished according to the internal relations of verbal and nominal elements (Kageyama 2009).

(9) a. Transitive VN + Object N:
   soo-kin 送金 [send-money] 'remit', *doku-syo* 読書 [read-book] 'read books', *sen-gan* 洗顔 [wash-face] 'wash one's face'

   b. Intransitive VN + Locative Complement:
   ki-koku 帰国 [return-country] 'return to one's country', *tai-sya* 退社 [leave-office] 'leave the office (for the day)'

   c. Unaccusative VN of appearance + Subject N:
   syuk-ka 出火 [go.out-fire] 'for a fire to break out', *raku-rai* 落雷 [fall-lightning] 'for lightning to strike (something)'

These are truly left-headed, since the left-hand elements originating from Chinese verbs and the right-hand ones deriving from Chinese nouns are arranged in the "V-N" order of Chinese syntax. In addition, the nominal and verbal morphemes that constitute them exhibit parallel distributions to native nouns and native verbs, respectively. For example, the nominal morpheme *kin* 'money' in *soo-kin* [send-money] 'remit' occurs in such S-J compounds as *syozi-kin* [possess-money] 'money carried in one's pocket' in the same way that the native noun *kane* 'money' occurs in *ari-gane* [have-money] 'all the money one has', and the verbal morpheme *soo* 'send' occurs in such S-J words as *soo-gei* [send-receive] 'welcoming and sending off' in the same way that the infinitive form of the native verb *okuru* 'send' occurs in the native compound *okuri-mukae* [send.off-receive].

The assumption that the verbal elements are the heads of the two-morpheme S-J words in (9) gains support from the strong correlation of transitivity between them and the entire compounds that are built on them. *Soo-kin* 'send money' is transitive as in *hyakuman-en o soo-kin(-suru)* [one-million-yen ACC send-money(-do)] 'remit one million yen' because *soo-* 'send' is transitive, and *ki-koku* 'return to one's country' is intransitive as in *nihon ni kikoku(-suru)* [Japan DAT return-country(-do)]

'return to Japan' because *ki-* 'return' is intransitive. Note that in conformity with the Chinese order of Adverbial + V, the order of nominal and verbal elements is reversed in the VNs whose noun members designate adjuncts of some kind or other, such as *doku-satu* [poison-kill] 'to kill with poison' and *zi-ritu* [by.oneself-stand] 'stand by oneself, become independent'. Again, the transitivity of such adjunct compounds is determined by their verbal elements: *doku-satu* [poison-kill] is transitive because *satu* 'kill' is transitive, and *zi-ritu* [by.oneself-stand] is intransitive because *ritu* 'stand' is intransitive.

Given that the compound VNs in (9) have their heads on the left, it can be legitimately conjectured that the left-hand verbal morphemes carry the lexical category of VN in themselves. In view of a general tendency that foreign words are borrowed as a kind of noun, it is highly conceivable that Chinese verbs took on a nominal character and were categorized as VNs when they entered the Japanese lexicon (Kageyama 1982). In fact, some S-J verbal morphemes have the liberty to function as independent words in Japanese syntax. For example, *datu-* 'escape' can be used as a verb when combined with *suru* 'do', as in *das-suru* 'escape from', suggesting that *datu-* itself is endowed with the category VN despite its bound status.

Even though the patterns of (9) are not fully productive, analogy with the existing words may give rise to new coinages from time to time, as illustrated by *bai-den* 売電 [sell-electricity] 'sell domestic solar power to a power company' and *bai-den* 買電 [buy-electricity] 'buy power', which became common after the power crisis caused by the disaster of the Fukushima nuclear power plant in 2011. New coinages may even involve non-S-J morphemes as their nominal elements, as in *datu-sara* 脱サラ [escape-salaried.worker] 'quit being an office worker', where *sara* is a clipping of *sararii-man* [salary-man] 'office worker', an English-like word made in Japan. Since this word functions as a VN in combination with the light verb *suru* 'do', *datu-* on the left is deemed the head. On the other hand, coinages like *datu-genpatu* 脱原発 [escape-nuclear.power.generation] 'quitting nuclear power generation at once' and *sotu-genpatu* 卒原発 [graduate-nuclear.power.generation] 'abandoning nuclear power generation step by step', whose second members comprise two morphemes/characters, are regarded as nouns rather than VNs, since they do not behave as verbs with *suru*. Presumably, *datu-* 'escape' and *sotu-* 'graduate' in these two examples are ad hoc prefixes created on analogy with *han-* 'anti-' in *han-genpatu* 反原発 [anti-nuclear.power.generation] 'anti-nuclear power generation', because they are pronounced with the phonological break characteristic of *han-* 'anti-' and other prefixes that are considered to make up a special morphological unit of "Word plus" (Kageyama 2001; Chapter 14 [Kageyama, this volume]).

Now, even a left-headed VN is subject to the regular right-headed structure when it is embedded in a larger compound, as in [[*soo-kin* (send-money)] + [*hoohoo* (method)]] 'the method of remittance' (Kageyama 2009). The shift of head position is common to all kinds of S-J compounds composed of three or more words, regardless of whether the larger compounds involve a predicate-argument relation or

a modifier-modified relation. For example, by using two two-character S-J words, *ken-poo* [law-law] 'constitution' and *kai-sei* [revise-right] 'revise', two larger compounds can be produced which differ in meaning, as in (10a) and (10b).

(10) a.   argument-predicate relation: [[*ken-poo*] [*kai-sei*]] (constitution revision) 'revision of the constitution'

      b.   modifier-modified relation: [[*kai-sei*] [*ken-poo*]] (revision constitution) 'revised constitution'

The different meanings in (10a) and (10b) correlate with their right-hand head structures.

The other kind of left-headed structure that has been frequently noted in the literature (Kageyama 1982) concerns Sino-Japanese negative prefixes including *mi-* 'not as yet', *mu-* 'no, null, non-existent', and *hu-* (NEG).

(11) a.   VN *seizyuku* 'ripen, become mature', as in *seizyuku-suru* [ripen-do] → AN *mi-seizyuku (na)* [not.as.yet-ripen (ENDING)] 'immature'

      b.   N *sekinin* 'responsibility' → AN *mu-sekinin (na)* [no-responsibility (ENDING)] 'irresponsible'

      c.   N *kigen* 'temper, mood' → AN *hu-kigen (na)* [NEG-temper (ENDING)] 'ill-tempered'

In (11a), a VN changes to an AN with the addition of the prefix *mi-*, and in (11b, c), Ns change to ANs by virtue of the prefixes *mu-* and *hu-*. These prefixes are thus generally assumed to have a category-changing function, as illustrated in (11). A radically different analysis will be discussed in Section 3.

2.3 Double-headed or coordinate compounds

Double- (or multiple-) headed compounds are coordinate compounds (also known as "dvandva") that consist of two or more members functioning as heads on an equal footing. There are a variety of combinations of lexical categories, as shown in (12).

(12) a.   N+N: *niti-bei* [Japan-U.S.A.] 'Japan and the U.S.A.

      b. $[V]_N+[V]_N$: *kasi-kari* [lending-borrowing] 'borrowing and lending'

      c.   V+V: *imi-kirau* [detest-hate] 'detest and hate'

      d.   A+A: *hoso-nagai* [thin-long] 'long and narrow'

      e.   AN+AN: *zyuukoo-tyoodai* [heavy.thick-long.big] 'big and heavy'

Referring to Wälchli (2005), Kageyama (2009) classifies these coordinate compounds into three types. The first is a holistic type, where the coordinated elements together refer to one conceptual unit, as in *oo-bei* lit. 'Europe-America' denoting 'the West'. The second is a relational type, where the compound indicates the notion of 'between X and Y', as in [*niti-bei*]-*kankei* 'Japan-U.S.A. relationship'. Particularly intriguing is the third type called a separate-reference type, where each of the coordinated elements has its own referent, as in *oya-ko* [parent-child] 'father/mother and/or children', *ten-ti* [sky-ground] 'heaven and earth', and *asa-ban* [morning-evening] 'morning and/or evening' (cf. also Nishio 1976; Nomura 1977). For example, *oya-ko* [parent-child] can be used with adverbs like *issyoni* 'together' that refer to a joint action of the participants as a group, or with adverbs like *sorezore* 'each' and *betubetuni* 'separately' that refer to separate actions taken by individual participants. This last type of coordinate compound has a characteristic accentuation in which the first member is pronounced with a high pitch and the subsequent members lose accent, as in *Oya-ko*, *TEn-ti*, and *Asa-ban* (capital letters stand for high pitches).

An interesting feature of coordinate compounds is that the order of members is constrained by cognitive and socio-cultural factors, in much the same way that the order of the English irreversible coordinates joined by *and* or *or* is fixed, as in *this and that* vs. **that and this* and *up and down* vs. **down and up* (Cooper and Ross 1975). Kageyama (1982) observes that the orders of Japanese coordinate compounds and English irreversible coordinates coincide in many cases but exhibit a systematic difference in certain semantic groups. The examples in (13) have the same order in the two languages. Each of these words is composed of two opposite or contrasting notions, one member of which is conceived of as having cognitive or socio-cultural priority over the other.

(13) a. Cognitive priority: *u-mu* 'existence or non-existence', *koo-hukoo* 'happiness or unhappiness', *sei-hi* 'right or wrong', *syuk-ketu* 'presence or absence', *yosi-asi* 'good or bad', *zen-aku* 'good and evil', *sin-gi* 'true or false', *kati-make* 'win or lose', *suki-kirai* 'like or dislike', *zyoo-ge* 'up and down'

   b. Socio-cultural priority: *dan-zyo* 'man and woman', *osu-mesu* 'male and female', *titi-haha* (native) 'father and mother', *hu-bo* (S-J) 'father and mother', *oya-ko* 'parent and child', *roo-nyaku* 'old and young', *syu-zyuu* 'master and servant', *si-tei* 'teacher and pupil', *koo-si* 'public and private', *zin-tiku* 'man and beast', *sei-kyoo* (politics-religion) 'church and state'

On the other hand, systematic differences are observed in deictic expressions, where English is centered on the speaker ('ego') and Japanese on non-speakers.

(14) *are-kore* [that-this] 'this and that', *ati-koti* [there-here] 'here and there', *yuki-ki* [go-come] 'come and go', *de-iri* [go.out-come.in] 'come in and go out', *dasi-ire* [take.out-take.in] 'take in and take out', *kasi-kari* [lending-borrowing] 'borrowing and lending', *uri-kai* [sell-buy] 'buy and sell'

Apart from these cognitive and socio-cultural factors, phonological factors appear to be at work in some cases. For example, both the native compound *migi-hidari* [right-left] and the Sino-Japanese compound *sa-yuu* [left-right] have a shorter morpheme in the first position.

    This section has shown that the headed word structures in Japanese are basically right-hand headed, but left-headed and double-headed structures are also observed marginally. It has also been suggested that the position of a categorial head is strongly associated with the semantic interpretation of a whole complex word. The next section will tackle headless or exocentric word structure.

3 Headless or exocentric word structure

It has long been assumed that headless or exocentric compounds are a minor, atypical kind of compound. Recently, however, Scalise, Fábregas, and Forza (2009) and Kageyama (2010) have brought to light their important role from typological and theoretical perspectives. From a typological perspective, Scalise, Fábregas, and Forza (2009), based on a database comprising some 3,000 compounds in 24 languages, present a relativized definition of exocentricity of complex words by differentiating three types of heads – categorial, semantic, and morphological – and propose to define three types of exocentricity accordingly. If we exclude the morphological head determined by features such as gender and number that are irrelevant to Japanese, we can confine ourselves to categorial and semantic exocentricity. This section will briefly review the arguments of Kageyama (2010), who provides a fairly comprehensive treatment of exocentric/headless compounds in Japanese by referring to both semantic and categorial factors.

3.1 Exocentric compounds due to figurative extension

In line with Benczes' (2006) work on English exocentric compounds, Kageyama (2010) maintains that the basic mechanisms of exocentric compound formation are metaphor and metonymy, which in Japanese are embodied in at least three different types: entity-to-entity mapping, event/state-to-entity mapping, and entity-to-event/activity mapping. Due to the figurative mappings, exocentric compounds convey special meanings that cannot be obtained by putting together the literal meanings of their constituents compositionally. In (15) are given some examples of entity-to-entity

mapping, by which a compound that literally designates a certain physical entity changes its meaning to represent a different type of entity.

(15) Entity-to-entity figurative mapping

    a. *kata-ude* [one-arm] 'one's right-hand man, one's reliable mate'

    b. *hiza-makura* [lap-pillow] 'someone's lap used as a pillow'

    c. *sizyuu-kata* [forty.years.of.age-shoulder] 'the kind of shoulder pain one often suffers from around forty years of age'

In (15a), the entire compound denoting a part of one's body ('one arm') is metonymically extended to designate one's reliable assistant or "right-hand man"; in (15b), the noun *hiza* 'lap' is metaphorically likened to the noun 'pillow'; in (15c), the compound that literally denotes the shoulders of a person around a particular age is extended to a typical physical condition arising in them. What is important is that the metaphorical and metonymic mappings of this kind are by no means a privilege of compound words but are commonly found with simplex nouns as well, as in *te* 'hand' being metonymically extended to mean 'helper, worker'.

The second type of figurative mapping illustrated in (16) is characteristically, though not exclusively, found with deverbal compound nouns consisting of a noun followed by a nominalized verb in the infinitive form.

(16) Event/state-to-entity metonymic mapping

    a. person: *e-kaki* [picture-drawing] 'painter', *kane-moti* [money-possessing] 'moneyed person', *nihon-zasi* [two.swords-wear] 'samurai, who typically wore two swords'

    b. entity: *ki-tutuki* [tree-peck] 'woodpecker (bird)', *waki-zasi* [side-wear] 'a short sword that a samurai wears beside a long one'

    c. instrument: *tume-kiri* [nail-cutting] 'nail clipper', *neko-irazu* [cat-need.not] 'rat poison'

    d. product: *tamago-yaki* [egg-frying] 'fried eggs, omelet', *musi-sasare* [bug-sting.PASS] 'bug bite'

    e. resultant state: *hada-are* [skin-getting rough] 'state of one's skin being rough'

    f. place: *mono-oki* [thing-keeping] 'storeroom'

    g. time: *yo-ake* [night-being.over] 'daybreak'

This kind of metonymic mapping is also ascribed to a general mechanism that applies not only to compound words but also to simplex deverbal nouns of the

native stratum, as shown by *suri* 'pickpocket' denoting a person (from the verb *suru* 'pick someone's pocket') and *hataki* 'duster' denoting an instrument (from the verb *hataku* 'to dust').

Unlike the first and the second types, the third type of entity-to-action metonymic mapping, shown in (17), appears pertinent only to compound nouns of the native stratum.

(17) Entity-to-action metonymic mapping
huka-zake [much-liquor] 'drink too much', *naga-buro* [long-bath] 'take a long bath', *huka-zume* [deep-nail] 'cut the nail to the quick', *tika-miti* [short-way] 'take a shortcut', *oki-tegami* [leave-note] 'leave a note'

The right-hand members of these words are all nouns denoting physical entities, but their actual meanings are shifted from the denotations of the right-hand members to the meanings of actions induced from the combined meanings of the two members. For example, *huka-zake* [much-liquor] does not mean a kind of liquor but the act of drinking too much liquor. Due to their action meanings, these compounds characteristically occur as the object of the verb *suru* 'do', as in *huka-zake o suru* [much-liquor ACC do] 'drink too much liquor'. Importantly, if the compound structure in the examples of (17) were disintegrated into its components, the kind of metonymic mapping would no longer obtain, as demonstrated by the ungrammaticality of **sake o suru* [liquor ACC do], where only the simple noun *sake* 'liquor' occurs, or by the bizarreness of **hukai sake o suru* [deep liquor ACC do], where the compound is paraphrased by a noun phrase.

The figurative extension observed with these special examples should be kept distinct from the regular polysemy of lexemes. For example, nouns like *isya* 'doctor' and *kyoosi* 'teacher' may be used to denote not only a particular human entity, as in *isya ga kita* [doctor NOM came] 'A doctor came', but also the particular profession he or she is engaged in, as in *isya o suru* [doctor ACC do] 'work as a doctor, practice medicine'. These nouns of professions will be conceived of as being lexically polysemous because no compounding process is involved. Given this, the entity-to-action metonymy in (17) is regarded as a special mechanism unique to compound nouns.

What would trigger the semantic shift in examples like those in (17)? Since the head nouns alone stand for physical entities, the elements in the modifier position should be responsible. For example, *huka-zake* [much-liquor] 'drink too much' is construed as meaning the act of drinking too much liquor because the modifier *huka-*, literally meaning 'deep' but here meaning an excessive amount of the liquor drunk, easily evokes the act of heavy drinking. Such a semantic shift is also attested in examples like those in (18), where an entity is reinterpreted as a state (Kageyama 2002).

(18) Entity-to-state metonymic mapping
han-doa [half-door] 'state of a car door being half open/not properly closed', *haya-mimi* [quick-ear] 'quick-eared, having a good ear for news'

An even more intriguing case is the beautific prefix *o-*, which, when prefixed to physical nouns, normally has the innocuous function of adding an elegant connotation to the nouns but exerts radical semantic effects when combined with certain nouns. Observe the examples in (19), where an entity is reinterpreted as an action or conversely, an action is reinterpreted as an entity (Kageyama 2002).

(19) a. entity-to-action mapping
te 'hand' → *o-te* 'a dog's action of shaking hands with a person'
hana 'flower' → *o-hana* 'Japanese-style flower arrangement'

      b. action-to-entity mapping
nigiri 'act of grasping' → *o-nigiri* 'a ball of rice'
mamori 'act of defending/protecting' → *o-mamori* 'good-luck charm'

The examples given in (17), (18), and (19) all involve semantic change triggered by the left-hand constituents of complex words, and such exocentric words are characteristically restricted to native vocabulary items. It is an interesting open question why complex words of the native stratum, which should be strictly right-hand headed in terms of morphological categories as we saw in Section 2, are susceptible to semantic shift under the influence of left-hand constituents.

3.2 Complex exocentric words due to form-meaning mismatch

While the figurative exocentric compounds discussed in Section 3.1 involve only semantic change, Kageyama (2010) draws attention to a novel type of exocentric words exemplified in (20), which appear to involve category change. These complex words, coming from both native and Sino-Japanese strata, share the morphological category of AN as seen from their compatibility with the ending *-na*.

(20) a. Native complex ANs of the form "adjective + noun"
hutop-para (na) [big-heart (ENDING)] 'big-hearted, generous', *oo-gara (na)* [big-physique (ENDING)] 'of large build', *yowa-ki (na)* [weak-mind (ENDING)] 'weak-minded', *oo-azi (na)* [rough-taste (ENDING)] 'lacking an exquisite taste'

      b. Sino-Japanese complex ANs of the form "prefix + noun"
ta-kinoo (na) [many-function (ENDING)] 'multi-functional', *aku-syumi (na)* [bad-taste (ENDING)] 'of vulgar taste', *koo-seinoo (na)* [high-performance (ENDING)] 'highly efficient', *ryoo-situ (na)* [good-quality (ENDING)] 'of good quality', *aku-situ (na)* [bad-quality (ENDING)] 'of poor quality', *i-situ (na)* [different-quality (ENDING)] 'heterogeneous', *koo-ka (na)* [high-price (ENDING)] 'expensive', *an-ka (na)* [low-price (ENDING)] 'cheap', *dai-kibo (na)* [big-scale (ENDING)] 'of large scale'

The peculiarity of these complex words is two-fold. Morphologically, despite the fact that their internal composition is entirely transparent, neither the left-hand nor the right-hand constituent exactly agrees with the AN category of the whole. The right-hand constituents, which are all nouns, do not take the *na* ending, and neither do the left-hand constituents. The left-hand elements in (20a) are the stems of native adjectives rather than full-fledged ANs, while those in (20b) are Sino-Japanese morphemes that look like prefixes and hence do not appear to have tangible morphological categories. Consequently, the complex words under discussion are not classified into any of the existing types of headed structures – right-headed, left-headed, or double-headed – in any straightforward way. In other words, they are headless or exocentric words with no discernible category heads inside them.

Semantically, these words are associated with peculiar interpretations that are not expected of a standard modification structure in which the first constituent modifies the second one. For example, *hutop-para* (lit. 'fat-belly') [big-heart] 'big-hearted, generous' does not refer to a physical entity that can be described as a protruded belly or a big heart, but instead designates a person's attribute of 'being big in heart', namely 'being big-hearted or generous'. According to Kageyama (2010), this interpretation obtains by paraphrasing the complex words in a subject-predicate relation with reversed order, as illustrated in (21a) with the native *hutop-para* and in (21b) with the Sino-Japanese *ta-kinoo*.

(21) a. *hutop-para* [big-heart] = *Hara ga hutoi.* [Heart NOM big] lit. 'one's heart is big' = 'generous' (*para* is a phonological variant of *hara* 'heart')

   b. *ta-kinoo* [many-function] = *Kinoo ga ooi.* [functions NOM many] lit. 'The functions are many' = 'multi-functional'

In (21), the meanings of the complex words are paraphrased with the corresponding native words by reversing the order of the first and second constituents into a subject-predicate relation. For example, *hutop-para* [big-heart] is plausibly paraphrasable as 'The heart is big', and *ta-kinoo* [many-function] as 'The functions are many'. Note that although the prefix-like constituents of the Sino-Japanese stratum in (20b) have no morphologically discernible category, they can be paraphrased naturally by employing their corresponding native adjectives in a manner similar to the *kanbun-kundoku* (漢文訓読) rendition where Chinese sentences are read by using native Japanese words à la Japanese grammar.

What is interesting is that the paraphrase formula in (21) exploits the so-called "double subject construction", a sentence structure characteristic of Japanese. As shown in (22), the double subject construction takes two nominative subjects, one of which, the small subject (Y), makes a composite predicate with the adjective, and this predicate phrase is in turn predicated of the big subject (X).

(22) [X *ga*   [Y *ga*   adjective]_{Predicate}].
      Syatyoo ga [president NOM]   *hara ga* [heart NOM]   *hutoi* [big]
      'The president is big-hearted (generous).'

Based on these morphological and semantic peculiarities, Kageyama (2010) proposes to attribute the nature of the exocentric words in (20) to a mismatch between morphology and semantics, as schematically shown in (23).

(23)   Morphology: A ('big') + Noun ('heart')

      Semantics: N ('heart') is A ('big').

As opposed to the figurative exocentric compounds whose meaning is opaque and unpredictable (Section 3.1), the exocentric words exemplified in (20) are transparent in both morphological composition and meaning, and their peculiarity is reduced to the twisted manner of their semantic interpretation. Kageyama (2010) claims that these exocentric words are composed of two dependents, i.e. Dependent (adjective) + Dependent (noun), and that the two dependent components collaborate to create a new hybrid category of Adjectival Noun only if the form-meaning mismatch in (23) arises.

    Kageyama (2010) further argues that the proposed analysis has the advantage of dispensing with the "category-changing" prefixes in Japanese. As noted in Section 2.2, the Sino-Japanese negative prefixes *mi-* 'not as yet', *mu-* 'no, non-existent, null', and *hu-* 'un-', have been regarded as exceptional prefixes that change the category of bases (N or VN) to AN. These S-J negative prefixes will no longer be exceptional but fall under the general schema of (23) if they are paraphrased by using the negative adjective *nai* in the native stratum, as exemplified in (24).

(24) *mu-sekinin*   = *Sekinin(kan)   ga   nai.*
    no-responsibility   responsibility NOM null
    'irresponsible'   'bear no responsibility'

There are in fact many more "category-changing" prefixes in the Sino-Japanese stratum, including *ta-* 'many', *koo-* 'high', and *dai-* 'big', as already shown in (20b). Although morphologically labeled as prefixes, these S-J morphemes as well as the three negative prefixes are considered to retain their original function as (adjectival) predicates in Chinese with their own lexical meanings. The analysis in (24) nicely captures the discrepancy between the morphological function of S-J bound morphemes as prefixes (left-hand elements) and their semantic function as adjectival predicates.

3.3 Syntactic effects of left-hand elements

This subsection will focus on cases where non-head elements in the left-hand position make non-trivial contributions to the syntactic usage of complex words. Roeper (1987) pointed out that the head constituent generally has the privilege of passing the subcategorization property on to a complex word in English, while a non-head constituent is prohibited from participating in the inheritance of subcategorization features, as demonstrated by the ungrammaticality of (25).

(25) a. *a <u>protection</u> plan **of** children (meaning 'a plan for protection of children')

    b. *a <u>tax</u>man **of** hidden assets (meaning 'a man who taxes hidden assets')

In (25), *of children* and *of hidden assets* are intended as the objects of *protection* and *tax* in the compound structures *protection plan* and *taxman*, respectively, but these interpretations are ruled out. Namiki (1985), however, observes some English examples where Roeper's constraint fails. In (26), the prepositions in boldface are felicitously associated with the underlined, non-head constituents of the complex words *guidebook* and *interconvertible*.

(26) a. a <u>guide</u>book **to** modern linguistics (cf. *a book to modern linguistics)

    b. Only the if-then type of condition is directly <u>inter</u>convertible **with** the rule format (cf. *convertible with the rule format)

    As far as English goes, this kind of phenomenon, which might be called "inheritance from non-heads", appears to be restricted to designated prefixes and particular compound nouns. However, it is found rather commonly with compound words in Japanese. Observe the examples of (27) pointed out by Sugioka (1989).

(27) a. ***tomodati e no* [*<u>tanomi</u>-goto*]**
        friend   DAT   GEN   ask-thing
        'a favor I asked my friend to do for me'

    b. ***minna kara no* [*<u>kiraw-are</u>-mono*]**
        everybody ABL   GEN   hate-PASS-person
        'person hated by everyone'

    c. ***tosyo no* [*<u>kasidasi</u>-kigen*]**
        book   GEN   lend-deadline
        'check-out limit for books'

    d. ***azia to no* [*<u>booeki</u>-koo*]**
        Asia   COM   GEN   trade-port
        'port for trade with Asian countries'

In (27), the noun phrases marked in the genitive are associated with the underlined left-hand constituents of the complex words. Thus in (27a), the noun phrase 'my friend' realizes the dative complement required by the verb *tanomi-* 'ask' in the

compound noun *tanomi-goto* 'something to ask' (cf. *tomodati ni tanomu* [friend DAT ask] 'ask my friend'). Likewise in (27d), the comitative phrase 'with Asia' represents the complement selected by the VN *booeki* 'trade' in *booeki-koo* 'trade port'. Moreover, inheritance from non-heads takes place regardless of the vocabulary strata of complex words: the head nouns in (27a, b) are native Japanese, while those in (27c, d) are Sino-Japanese. Unlike English, Japanese allows this kind of inheritance from non-heads fairly freely. Sugioka (1989) accounts for the disparity between English (26) and Japanese (27) by the contiguity of external phrases and non-heads in linear order. In Japanese, external phrases are adjacent to non-heads, both appearing to the left of the heads of complex words, while in English the relevant elements are not contiguous in linear order. Non-contiguous modification disrupts inheritance, as Sugioka assumes. See Chapter 14 (Kageyama, this volume) for more examples of inheritance from non-head positions.

An analogous phenomenon is observed with Verb-Verb compound verbs as well (Kageyama 1993; Yumoto 2005). The general principle of argument realization in compound verb constructions is that the verb in the second position (V2) of a compound determines the subcategorization of a whole compound, as exemplified by the compound verb *arai-nagasu* [wash-let.flow] 'wash away' in (28).

(28) *Watasi   wa   kuruma   no   **yogore**   o   [arai-**nagasi**-ta]*.
      I   TOP   car   GEN   dirt   ACC   wash- let.flow-PST
      'I washed away the dirt on the car.'

The verb *arai-* 'wash', the left-hand constituent of the compound verb, is associated with the noun 'car' on which the subject splashes water, whereas the verb *nagasu* 'let flow' in the head position is linked to the noun 'dirt' which flows away. In the syntactic structure of (28), the whole compound takes the noun *yogore* 'dirt' as the object marked in the accusative. Were the noun 'car' put in the accusative object, the sentence would mean that the car itself was washed away.

Contrary to this general principle of inheritance, Yumoto (2005) noticed that in some compound verbs, the verb in the first position (V1) is responsible for the selection of complements in a whole compound verb. Observe the examples in (29).

(29) a. *Watasitati   wa   onazi   **densya**   ni/*o   [nori-awase-ta]*.
          we   TOP   same   train   DAT/*ACC   get.on-happen-PST
          'We happened to get on the same train.'

      b. *Kanozyo   wa   **koibito**   **o**   [mati-kuras-ita]*.
          she   TOP   boyfriend   ACC   await-live-PST
          'She waited for her boyfriend for a whole day/many days.'

      c. *Titioya   wa   **musuko**   **o**   [sikari-tuke-ta]*.
          father   TOP   son   ACC   scold-do.violently-PST
          'Father scolded his son thoroughly.'

In (29a), the dative marking on the noun 'train' originates from the V1 *nori-* 'get on' of the compound verb *nori-awaseru* 'happen to ride', whereas the V2 *awaseru* lit. 'put together' has lost its original meaning in this compound and means that the multiple actions denoted by the verb in V1 took place coincidentally. Similarly, in (29b), the accusative object 'her boyfriend' is selected by the V1 *mati-* 'await', while the V2 *kurasu* lit. 'live, pass a day' means here only that the waiting activity continued for a long stretch of time; in (29c), the accusative object 'the son' is selected by the V1 *sikari-* 'scold', with the V2 *tukeru* emphasizing the severity of the scolding action. In these V-V compound verbs, the left-hand members play a vital role in the realization of argument structure and case.

Where does the difference between (28) and (29) come from? In Chapter 8 (Kageyama, this volume), it is shown that the verbs in the right-hand position of the (29)-type V-V compounds are devoid of argument structure and case and instead supply the verbs on the left with a variety of aspectual meanings. In such "aspectual compound verbs", the subcategorization features of a whole compound verb are regulated by the verbs in the non-head position.

To summarize, this section has focused on cases in which the left-hand member of a complex word plays a pivotal role in determining the category, meaning, and subcategorization of the whole. These cases illustrate the complexity of head relations in Japanese morphology, which goes beyond the simple rule of right-hand head.

4 On the boundary between compounding and affixation

The discussion in the foregoing sections concerning categorial heads and the semantic effects of non-heads now leads us to reconsider the traditional distinction between compounding and affixation: compounding combines two (or more) words to form another word, while affixation attaches a prefix or suffix to a word to form another word. Although we do not intend to deny the basic validity of the distinction, cases are occasionally encountered in which it is hard to determine whether a given morpheme is a compounding element or an affix resulting from semantic bleaching or grammaticalization.

4.1 Semantic bleaching in the left-hand constituents of compounds

Namiki (2013) presents interesting examples like those in (30) where a noun takes on a new meaning that departs from the original when used in the left-hand position of compounds.

(30) a. *mame* 'bean' : *mame-denkyuu* [bean-light.bulb] 'very small light bulb', *mame-taihuu* [bean-typhoon] 'very small typhoon', *mame-tisiki* [bean-knowledge] 'a small piece of knowledge'

    b. *hime* 'princess': *hime-kyoodai* [princess-dressing.stand] 'small dresser', *hime-yuri* [princess-lily] 'small lily'

    c. *oni* 'ogre' : *oni-hitode* [ogre-starfish] 'large starfish, crown-of-thorns starfish', *oni-yanma* [ogre-dragonfly] 'jumbo dragonfly'

    d. *kusa* 'grass' : *kusa-keiba* [grass-horse.race] 'local horse race', *kusa-yakyuu* [grass-baseball] 'sandlot baseball'

In (30a), the noun *mame* meaning 'bean' or 'pea' conveys the meaning of 'small' when used as the left-hand member of a compound. Similarly, *hime* in (30b), which lexically denotes 'princess' in its independent use, expresses the meaning of 'small and cute' when used in the modifier position of some compounds (the semantic shift of *hime* was already attested in *hime-gaki* 'small hedge' in Early Middle Japanese (Sakakura 1966: 427)). Analogous semantic shifts are observed in (30c) and (30d) with the nouns *oni* 'ogre' and *kusa* 'grass', which respectively convey the meanings of 'large and unusual' and 'local or nonprofessional' when they appear in the modifier position of these compounds. These semantic shifts are undoubtedly attributed to metaphor or metonymy: the transferred meaning of 'small' in *mame*, for example, derives from the physical smallness of beans.

Although such figurative extension of meaning is in itself quite common in the world's languages, the examples reported in the literature are mostly concerned with simplex words or compound words as a whole (cf. the discussion on exocentric compounds in Section 3.1). What is vitally important about examples of (30)-type is that the semantic bleaching (or "dilution of meaning" in Namiki's terminology) takes place solely in the first, modifier position of certain designated compounds, and never occurs in the second, head position of any compound. Compare thus the compounds in (30a) with *endoo-mame* 'green pea' and *nankin-mame* [Chinese-bean] 'peanut', or those in (30b) with *sirayuki-hime* [white.snow-princess] 'Snow White', where the nouns in question maintain their literal meanings when used in the head position. Here one might point out that the noun *mammoth* in English exhibits a similar behavior in expressions like *mammoth redwood*. However, dictionaries give the category of Adjective to this use of *mammoth* meaning 'huge'. Such relabeling is inconceivable for the examples of (30), which are genuine compound words rather than syntactic noun phrases. Booij (2010) adduces analogous examples from Dutch, proposing the notion of "construction-specific meaning" (meaning that is specific to certain morphological constructions). The Japanese examples in (30) will also be subsumed under this category.

Similar phenomena are observed with Verb-Verb compound verbs as well. Take *buti-* in *buti-kiru* [strike-cut] 'hack' and *buti-kireru* [strike-be.cut] 'flip out (in a rage)' for example. Here the first verb *buti-* (dictionary form *butu* 'strike, hit hard') has lost its physical meaning and changed to a more abstract meaning of emphasizing the hardness or forcibleness of an action. This semantic bleaching is correlated with the phonetic erosion where *buti-* is realized in some cases as *but-* with the drop of the vowel *i*, as in *but-tigiru* 'tear forcibly' and *buk-korosu* 'kill forcibly'. Because of this, dictionaries give *buti-* or *but-* as a prefix rather than as a full-fledged verb (for more examples of this sort, see Chapter 8 [Kageyama, this volume]).

In the literature on cognitive linguistics, proposals are made to distinguish several stages through which grammaticalization proceeds, although changes do not have to go through all the stages but may stop at a certain stage (Hopper and Traugott 2003: 131). For example, Heine and Kuteva (2002) disintegrate the broad notion of "grammaticalization" into "desemanticization" (semantic bleaching) as the first stage, "decategorization" (category change) as the second stage, and phonetic erosion (change in phonetic shape) as the last stage. Although both the *buti-* phenomenon and the compound examples in (30) fall in the broad coverage of grammaticalization, they are at different stages of grammaticalization. The compounds in (30) remain at the stage of desemanticization, whereas the *buti-* examples have proceeded to the next stage of decategorization.

4.2 Semantic bleaching in the right-hand constituents of compounds

The weakening of meaning is observed with the second constituents of some compounds as well. Illustrative examples are given in (31), many of which are new and unestablished coinages (Namiki 2009).

(31) a. *onti* 'tone-deaf, poor ear for music': *hookoo-onti* [direction-tone.deaf] 'poor sense of direction', *undoo-onti* [exercise-tone.deaf] 'poor exercise capacity'

     b. *nanmin* 'refugee': *kaigo-nanmin* [nursing.care-refugee] 'senior citizen who has difficulty in finding a nursing home', *osan-nanmin* [childbirth-refugee] 'an expecting woman who has much difficulty in finding a hospital for delivery'

     c. *somurie* 'sommelier (Fr.)': *yasai-somurie* (vegetable-sommelier) 'expert in selecting vegetables', *taoru-somurie* (towel-sommelier) 'expert in selecting towels'

The noun *onti* in (31a) means 'tone-deaf' or 'having no ear for music, inability to sing a song well' when it is used as an independent noun or as the left-hand constituent

of a compound as in *onti-kyoosei* 'improvement of the singing ability'. However, when it occurs in the right-hand position of a compound, it waives its specific reference to 'auditory or musical ability' and applies to a variety of human activities that call for an agent's skill. Thus, *hookoo-onti* [direction-poor.ability] refers to an agent's poor sense of direction, and *undoo-onti* [exercise-poor.ability] to an agent's poor ability in exercise or sports. Likewise in (31b), the noun *nanmin*, by itself designating 'refugee' or 'one who is unwillingly displaced from his/her country', refers to one who is displaced from, or cannot find his or her place in, an institution where he/she can settle when it is used as the head of a compound, as in *osan-nanmin* [childbirth-refugee] 'a pregnant woman who is not accommodated in a hospital for delivery although she wants to'. When used as the head of a compound, the noun *somurie* 'sommelier' in (31c) generalizes its original specification of wine selection to the ability of evaluating any kind of artifact, as in *yasai-somurie* [vegetable-sommelier] 'an expert at choosing vegetables' and *taoru-somurie* [towel-sommelier] 'an expert at selecting towels', although these novel coinages sound facetious at present. The semantic bleaching observed in these examples take place regardless of the vocabulary strata, as seen from the fact that *yasai* 'vegetable' is Sino-Japanese and *somurie* is foreign.

While all the examples discussed in this section are common nouns, proper nouns, especially place names, are also susceptible to semantic bleaching. Namiki (2009) provides coinages like those in (32).

(32) a. *Koosien* (Japanese ballpark in Koshien; famous for high school baseball championships), metonymically, 'high school baseball championships held at this stadium': *haiku-koosien* [haiku.poem-Koshien] 'a high school championship event for making *haiku* poetry', *syasin-koosien* [photograph-Koshien] 'a high school photography competition'

  b. *Ginza* (a popular business and shopping center of Tokyo): *Togoshi-ginza* 'name of the shopping street in Togoshi, Shinagawa City', *taihuu-ginza* [typhoon-Ginza] 'places frequently struck by typhoons'

In (32a), *Koosien* (Stadium), a ballpark in Japan that is famous as the venue for high school baseball championships, has lost its original geographical referent and is construed as a common noun designating the venue for certain competitive activities involving high school students. This is a case of metonymic transfer. In (32b), on the other hand, Ginza, one of the busiest shopping and business towns in Tokyo, is used to name the bustling shopping streets in rural towns by metaphor. Street names suffixed with *ginza* are fairly common throughout Japan. In contrast, *taihuu-ginza* [typhoon-Ginza] is not a street name but refers to 'a place that is frequently attacked by typhoons' – a cynical naming that likens unwelcome typhoons to welcomed shoppers and visitors. In a nutshell, the geographical locations of Koshien and Ginza

are no longer relevant in (32), and only the noteworthy features pragmatically associated with them are deployed in the figurative meanings of the heads in the compounds.

The examples presented in this subsection all share semantically bleached head nouns. Morphologically, such expressions will be analyzed as compounds because the head nouns can be used independently though with slightly different meanings. Their positional restriction suggests, however, that these head nouns with bleached meanings might be better viewed as suffixes rather than compounding elements.

4.3 From words to suffixes

The tendency for compounding elements to change to suffixes with dilution of their original meanings is observed with categories other than nouns as well. *Itiryuu* 'first-rate', for example, normally functions as a prenominal modifier in noun phrases like *itiryuu no sikisya* [first-rate GEN conductor] 'eminent conductor' and in compound nouns like *itiryuu-hoteru* [first.class-hotel]. Namiki (2009) observes, however, that the same word acquires an entirely different meaning of 'characteristic (of)' or 'peculiar (to)' when it occurs as the second constituent of a complex word that takes the form of N-*itiryuu (no)* [N-first.class GEN], as exemplified by (33a).

(33) a. [*Sanuki-huzin-itiryuu*]   *no*   *kokoro* *nimo* *nai*
        Sanuki-woman-specific GEN mind   in   absent

        kenson   *no*   *poozu*
        modest   GEN   pose
        'modest behavior characteristic of Sanuki women'

   b. [*syuugakuryokoosei-yorosiku*]   *osyaberi* *ni*   *mutyuu*   *ni*   *nari,*
        school.excursion.student-just.like chatting DAT absorbed DAT get
        '(We) got absorbed in chatting just like students on school excursion'

An even more interesting case is (33b). The word *yorosiku*, the *ren'yō* form of the adjective *yorosi-i* 'good, fine', is commonly used as a fixed expression in asking someone a favor (*Yorosiku onegaisimasu* 'I hope you'll take good care of this') or giving one's regards to someone (*Okusama ni yorosiku* 'Give my best regards to your wife'). When used in the morphological construction "N-*yorosiku*" as in (33b), however, the word loses its adjectival or adverbial meaning altogether and takes on the meaning 'just like N' or 'in the manner of N'. What is important is that this particular meaning is strictly limited to the right-hand position of the pattern N-*yorosiku*. This structural restriction, coupled with the special meaning tightly linked with the morphological construction, strongly indicates that -*yorosiku* as well as -*itiryuu* is grammaticalized to a suffix with a construction-specific meaning.

The assumption that *-itiryuu* and *-yorosiku* behave as suffixes in the particular usage gains support from the following observations. First, N-*itiryuu* looks at first glance like a compound noun similar to the compounds shown in brackets in (34).

(34) a. [*Koobe-tokuyuu*]   *no*   *sakamiti*
          Kobe-specific   GEN   uphill
          'uphills characteristic of Kobe'

      b. [*sumiyaki-gama-dokutoku*]   *no*   *nioi*
          charcoal.burning-kiln-particular   GEN   scent
          'the scent particular to a charcoal kiln'

A characteristic property of the compounds in (34) is that they are paraphrased by using the dative particle on the first members, as shown by (35a) and (35b).

(35) a. *Koobe* **ni**   *tokuyuu*   *no*   *sakamiti*
          Kobe   DAT   characteristic   GEN   uphill
          'uphills that are characteristic of Koobe'

      b. *sumiyaki-gama*   **ni**   *dokutoku*   *no*   *nioi*
          charcoal.burning-kiln   DAT   particular   GEN   scent
          'the scent that is particular to charcoal-making kilns'

Such a paraphrase with a dative-marked complement, however, is unavailable for N-*itiryuu*, as demonstrated by the ungrammaticality of (36).

(36) *Sanuki-huzin   *ni*   *itiryuu*   *no* . . .   cf. (33a)
          Sanuki-women   DAT   characteristic   GEN

The paraphrase with a dative complement is also rejected by N-*yorosiku*. The complex word in (33b) is not synonymous with the phrasal paraphrase in (37).

(37) **syuugakuryokoosei*   *ni*   *yorosiku* . . .   cf. (33b)
          school.excursion.student   DAT   just like

(37) makes sense only if it is intended to convey the literal meaning of 'Say hello to the students on school excursion'.

We have thus shown that *itiryuu* and *yorosiku* convey special meanings only when they appear in the second position of complex words. The semantic and positional restrictions plausibly suggest that they are not so much compounding elements as suffixes. It is notable that contrary to the standard view that semantic bleaching generally originates in metaphor or metonymy, their specialized meanings

cannot be considered as figurative extensions of their original meanings. While Namiki (2001, 2009) calls such special meanings "compound-specific submeanings," it will be more fitting to subsume them under a broader notion of "construction-specific meanings" in the sense of Booij's (2010) Construction Morphology because *itiryuu* and *yorosiku* appear to have attained a suffixal status.

An example of a similar though not exactly the same kind is the bound morpheme *-hoodai* 'without restriction, freely', which is listed in dictionaries as a suffix attaching to verb infinitives or to some other words like *katte* 'free will' and V-*tai* 'want to V'. In earlier Japanese, it was used as an independent noun designating a volitional act of behaving at one's will (without restriction), and this meaning is inherited by the contemporary suffix *-hoodai*, as illustrated by *iitai-hoodai* [say.want-without.restriction] 'saying what one feels like saying', *tabe-hoodai* [eat-without.restriction] 'eating as much as one wants', and *tirakari-hoodai* [be.messy-without.restriction] 'being hopelessly messy and untidy (said of a room or a place)' (Namiki 2010). Because of its positional restriction, *hoodai* cannot appear as the first constituent of a complex word, as shown by the total ungrammaticality of **hoodai-iitai* [*hoodai*-want.say] and **hoodai-tirakari* [*hoodai*-be.messy].

What is unique about *-hoodai* is that it can be attached to syntactically motivated elements such as causative *-sase*, passive *-rare*, light verb constructions, and idioms in the same way that the head verbs of syntactic V-V compound verbs can (see Chapter 8 [Kageyama, this volume]). Observe the examples in (38).

(38) a. causative: *tabe-sase-hoodai* [eat-CAUS-without.restriction]

  b. passive: *sawar-rare-hoodai* [touch-PASS-without.restriction]

  c. light verb construction: *tootyoo-si-hoodai* [wiretap-do-without.restriction]

  d. idiom: *abura o uri-hoodai* [loaf.around-without.restriction]

The compatibility with such syntactic elements indicates that *hoodai* is attached in syntactic structure rather than being affixed to the base noun directly in the lexicon. Strong evidence supporting this analysis is found in the occurrence of the accusative case marker on the arguments of the verbs it attaches to, as shown in actually occurring examples in (39).

(39) a. *syasin   o   tori-hoodai*
       photo   ACC   take-without.restriction
       'taking pictures without restriction'

  b. *denwa   o   tootyoo-s-are-hoodai*
       phone   ACC   wire.tap-do-PASS-without.restriction
       'letting the phone to be wiretapped freely'

In (39), *hoodai* behaves on a par with syntactic auxiliaries like *-sooda* 'it looks like', which also maintain the case marking of the arguments of their host verbs.

On the other hand, the examples in (40) suggest that *hoodai* has a nominal function as well, because the arguments of the host verbs are realized by the genitive instead of accusative case.

(40) a. *biiru   no   nomi-hoodai*
          beer   GEN   drink-without.restriction
          'drinking beer without restriction'

      b. *denwa   no   kake-hoodai*
          phone   GEN   call-without.restriction
          'making phone calls without restriction'

Interestingly, this nominal use appears to be incompatible with syntactic elements like causative and passive.

(41) a. **karada   no   sawar-rare-hoodai*
          body   GEN   touch-PASS-hoodai
          'being touched on the body without restriction'

      b. **denwa   no   tootyoo-s-are-hoodai*
          phone   GEN   wire.tap-do-PASS-without.restriction
          'being tapped without restriction'

These observations show that *hoodai*, originally an independent noun, has developed into two different kinds of bound morpheme, one as a syntactic auxiliary (39) and the other as a lexical suffix (40). It can be conjectured that the original meaning of *hoodai* denoting 'a volitional act of behaving at one's will' was shifted to an adverbial meaning of 'freely' or 'without restriction' and further bleached to an aspectual meaning of 'continuation without an explicit end'.

Despite the decategorization as a noun-forming suffix or a nominal auxiliary, however, *-hoodai* retains its original phonological properties, as indicated by the compound accent it exhibits (Namiki 2010). The compound accent is a phonological indicator of two words being integrated into one morphological unit with a single accent peak. Observe (42), where capital letters indicate high pitches (cf. Kubozono 1995).

(42) a. *SOosa*   + *HOnbu*   → *soOSA-HOnbu*
         investigation   headquarters   investigation-headquarters

     b. *gaikokugo*   + *gakubu*   → *gaIKOKUGO-GAkubu*
         foreign.language   department   foreign.language-department

The form "X-*hoodai*" also follows the compound accent pattern, as shown in (43).

(43) *tabe* + *HOodai*   → *taBE-HOodai*
       eat   without.restriction   eat-without.restriction
       'eating as much as one wants'

It is thus concluded that the grammaticalization of *-hoodai* has not proceeded to the last stage of phonetic erosion.

4.4 From affixes to words

In early theories of grammaticalization, the hypothesis was entertained that grammaticalization processes take place unidirectionally, whereby full-fledged words may be grammaticalized to bound affixes but the opposite direction of affixes becoming independent words is explicitly denied. Hopper and Traugott (2003), however, caution against taking the hypothesis as an absolute: despite the overwhelming evidence in its support, counterexamples are sporadically observed. A well-known example of "degrammaticalization" is the possessive clitic (e.g. English *'s*) that grew out of the genitive suffix in some Germanic languages (cf. Norde 2009; Traugott and Trousdale 2013). In the literature on Japanese grammar, two verbs are often cited as cases of degrammaticalization in which affixes grew into independent words.

One such case is *-mekasu*, which may be used both as a verb-forming suffix with the meaning of 'behave like, pretend to be, try to show', as in (44), and as an independent verb with the meaning of 'dress up', as in (45).

(44) a. *sinsetu-mekasu* [kind-SUF] 'pretend to be a kind person'

      b. *zyoodan-mekasu* [joke-SUF] 'say jokingly'

(45) a. *Kanozyo   wa   mekasi-te   deeto   ni   it-ta.*
          she   TOP dress.up-GER date   to   go-PST
          'She dressed up and went to a date.'

      b. *mekasi-komu* (V-V compound verb)
          dress.up-do.fully 'dress up fully'

Semantically, these two usages are related with each other by the meaning of 'outward display of a concrete or abstract attribute'. According to the unidirectionality hypothesis, the usage of *mekasu* as an independent verb (45) should constitute the basis from which the affixal usage in (44) emerged. Historical evidence shows, however, that the contrary was true.

According to Sakakura (1966: 137), Old and Middle Japanese exhibit only the suffixal use of *-mekasu* meaning 'to make exhibit a characteristic property', as in (46).

(46)　*hito-mekasu* [person-SUF] 'treat as a decent person', *ima-mekasu* [now-SUF]
　　　'make look modern', *iro-mekasu* [color-SUF] 'make showy', *mono-mekasu*
　　　[thing-SUF] 'make stand out'

Sakakura (1966) further contends that this suffix later developed into a full-fledged verb with the meaning of 'to dress up', as shown in (45a) above – a claim supported by historical documentation in which the earliest appearance of the suffixal *-mekasu* dates back to c. 1000 whereas the independent use first occurred in 1801 (*Nihon kokugo daijiten* 2000–2002). Murayama (2014: 136) also adduces several examples of *mekasu* used as an independent word in 1801–1835, but not earlier. Norde (2009: 8) claims that "there are no examples of degrammaticalization 'all the way up the cline' – a degrammaticalization chain from suffix all the way to lexical item has not been attested." If the historical development of *mekasu* from a suffix to an independent verb is correct, however, it would count as a counterexample to Norde's generalization.

　　Sakakura makes an additional suggestion that the verb-forming suffix *-buru* 'behave as if, pretend to be', used as in (47a), and the verb *buru* 'put on airs', used as in (47b), might be another candidate of change from a suffix to an independent word.

(47)　a.　*gakusha-buru* [scholar-pretend] 'act pedantically', *zyoohin-buru* [decent-pretend] 'give oneself airs', *era-buru* [important-pretend] 'act pompously'

　　　b.　*Sonnani*　*bur-anaku-te*　　　　*ii*　　*yo.*
　　　　　so.much　put.on.airs-NEG-GRD　fine　SFP
　　　　　'You don't have to put on airs like that.'

In fact, the earliest attested example of the suffix *-buru* goes back to c. 1485, and that of the verb *buru* to 1841 (*Nihon kokugo daijiten* 2000–2002, s.v. *buru*). In both the *mekasu* and *buru* cases, the semantic change followed the path of particularization, which is opposite to semantic bleaching of the familiar kind.

　　To summarize, this section has presented Japanese examples that appear relevant to the fundamental division of word formation processes into compounding and affixation. This distinction, albeit basically tenable, was called into question by considering certain compounding elements that behave like prefixes when they are used as the first constituents of compounds (e.g. *hime* 'princess, small') or like suffixes when they are used as the second constituents of compounds (e.g. *hoodai* 'without restriction'). In the last subsection, the behaviors of *mekasu* and *buru* were discussed as potential counterexamples to the unidirectionality hypothesis of grammaticalization.

5 Conclusion and future research perspectives

This chapter has presented both universally endorsed phenomena and language-particular phenomena commonly observed in Japanese. In constructing a universally valid model of grammar, the language-particular phenomena should be of special interest. The discussion on Japanese data, especially on the interaction of morphemes coming from two distinct lexical strata, Sino-Japanese and native Japanese, presents a conclusion that should be challenging to existing theories of headedness and grammaticalization built in a large measure on European languages. Concerning the notion "head of a word", Williams' (1981) Right-hand Head Rule alone is incapable of accounting for the wide spectrum of headed and headless word structures available in Japanese. The characterization of a head itself is also problematic. In this chapter, we have intentionally focused on the categorial property as the defining criterion of a head without committing ourselves to hybrid theories that allow reference to both categorial and semantic properties, as in Aronoff and Fudeman (2005), or distinguish three kinds of head – categorial, semantic, and morphological – as in Scalise, Fábregas, and Forza (2009). The Japanese data discussed in this chapter suggest that semantic and morphological (agreement) information should be separated from the notion of categorial head. Although categorial head and semantic center converge in many cases, they are occasionally realized in different positions of a word, as in Japanese cases where left-hand or non-head constituents exert non-trivial semantic effects on the meaning of whole complex words. Although the notion of categorial head is fairly robust, lexical semantic information will have to be disintegrated into such elements as conceptual meaning and subcategorization features. It is also necessary to elucidate the syntactic role of non-head elements in determining the argument and case relations in sentences in which the complex words are used.

Related to the headedness issue is the boundary between compounding and affixation, which tends to be blurred in Japanese. A few selected factors can be brought to bear upon this issue. One is the morphological character of Japanese as an agglutinative language, where both compounding elements and suffixes/auxiliaries are iteratively added to the right of a base. The other is the peculiarity of Sino-Japanese affixes, which have a variety of lexical meanings almost on a par with independent words (see Chapter 3 [Kobayashi, Yamashita, and Kageyama, this volume] for Sino-Japanese affixes). Of utmost importance will be the degree to which the meaning of an element correlates with its morphological position. As demonstrated by the examples given in Section 4, it may happen that compounding elements are fixed in the position they can occupy in complex words and are concomitantly limited to specific meanings. These elements will then be looked upon as being intermediate between words and affixes. Any theory of compounding and affixation must pay due attention to such intermediate cases. In Section 4, we also touched on the issue of directionality in grammaticalization, observing two sets of cases, one in which words changed to affixes and one in which affixes grew into words.

Acknowledgments

References

Allen, Margaret. 1978. *Morphological investigations*. Storrs, CT: University of Connecticut dissertation.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark and Kirsten Fudeman. 2005. *What is morphology?* Malden, MA: Blackwell.

Bauer, Laurie. 1990. Be-heading the word. *Journal of Linguistics* 26. 1–31.

Benczes, Reka. 2006. *Creative compounding in English: The semantics of metaphorical and metonymical noun noun combinations*. Amsterdam: John Benjamins.

Bloomfield, Leonard. 1933. *Language*. London: George Allen and Unwin.

Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.

Booij, Geert. 2012. *The grammar of words*. 3rd edition. Oxford: Oxford University Press.

Cooper, William and John Ross. 1975. World order. In Robin E. Grossman, L. James San and Timothy J. Vance (eds.), *Papers from the Parasession on Functionalism*, 63–111. Chicago: Chicago Linguistic Society.

Di Sciullo, Anna Maria. 2005. *Asymmetry in morphology*. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 2001. Word plus: The intersection of words and phrases. In Jeroen van de Weijer and Tetsuo Nishihara (eds.) *Issues in Japanese phonology and morphology*, 245–276. Berlin: Mouton de Gruyter.

Kageyama, Taro. 2002. *Kejime no nai nihongo* [Japanese as an unbounded language]. Tokyo: Iwanami Shoten.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford handbook of compounding*. 512–526. Oxford: Oxford University Press.

Kageyama, Taro. 2010. Variation between endocentric and exocentric word structures. *Lingua* 120. 2405–2423.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kobayashi, Hideki, Kiyo Yamashita and Taro Kageyama. this volume. Chapter 3: Sino-Japanese words. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kubozono, Haruo. 1995. *Gokeisei to on'inkōzō* [Word formation and phonological structure]. Tokyo: Kurosio Publishers.

Murayama, Miwako. 2014. Dōshi "mekasu" no seiritsu [The rise of the verb *mekasu*]. *Kokugo Goishi no Kenkyū* [Studies on the history of Japanese vocabulary] 33. 129–145. Osaka: Izumi Shoin.

Nagano, Akiko and Masaharu Shimada. 2014. Morphological theory and orthography: *Kanji* as a representation of lexemes. *Journal of Linguistics* 50. 323–364.

Namiki, Takayasu. 1985. *Gokeisei* [Word formation]. Tokyo: Taishukan.

Namiki, Takayasu. 1994. Heads and subheads of compounds. In Shuji Chiba, Yasuaki Fujiwara, Masayuki Ike-uchi, Tsuguyo Kohno, Osamu Koma, Yukio Nagahara, Takayasu Namiki, Yukio Otsu, Masanori Suiko, Kazuhiro Ushie and Norio Yamada (eds.), *Synchronic and diachronic approaches to language: A festschrift for Toshio Nakao on the occasion of his sixtieth birthday*, 269–285. Tokyo: Liber Press.

Namiki, Takayasu. 2001. Further evidence in support of the right-hand head rule in Japanese. In Jeroen van de Weijer and Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*, 277–297. Berlin and New York: Mouton de Gruyter.

Namiki, Takayasu. 2009. *Tango no kōzō no himitsu: Nichieigo no zōgohō o saguru* [The secrets of the structure of words: inquiry into word formation of English and Japanese]. Tokyo: Kaitakusha.

Namiki, Takayasu. 2010. Morphological variation in Japanese compounds: The case of *hoodai* and the notion of "compound-specific submeaning." *Lingua* 120. 2367–2387.

Namiki, Takayasu. 2013. Fukugōgo to haseigo [Compounds and derivatives]. In Taro Kageyama (ed.), *Lexicon forum* 6, 43–57. Tokyo: Hituzi Syobo.

Nihon kokugo daijiten [The grand dictionary of the Japanese language]. 2000–2002. 2nd edition. Tokyo: Shōgakukan.

Nishio, Toraya. 1976. Zōgohō to ryakugohō [Word building and word abbreviation]. In Suzuki Takao (ed.), *Nihongo no goi to hyōgen* [Vocabulary and Expressions of Japanese], 27–62. Tokyo: Taishukan Shoten.

Nomura, Masaaki. 1977. *Zōgohō* [Word building]. In Susumu Ōno and Takesi Sibata (eds.), *Nihongo 9: Goi to imi* [Japanese 9: Vocabulary and meaning], 245–284. Tokyo: Iwanami Shoten.

Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.

Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18. 267–310.

Sakakura, Atsuyoshi. 1966. *Gokōsei no kenkyū* [A study of word composition]. Tokyo: Kadokawa Shoten.

Scalise, Sergio. 1992. Compounding in Italian. *Rivista di Linguistica* 4. 175–199.

Scalise, Sergio, Antonio Fábregas and Francesca Forza. 2009. Exocentricity in compounding. *Gengo Kenkyu* 135. 49–84.

Selkirk, Elisabeth. 1982. *The syntax of words*. Cambridge, MA: MIT Press.

Sugioka, Yoko. 1989. Haseigo ni okeru dōshi-sosei no uketsugi [Inheritance of verbal features in derived words]. In Susumu Kuno and Masayoshi Shibatani (eds.), *Nihongogaku no shintenkai* [New developments of Japanese linguistics], 167–185. Tokyo: Kurosio Publishers.

Traugott, Elizabeth Closs and Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Wälchli, Bernhard. 2005. *Co compounds and natural coordination*. Oxford: Oxford University Press.

Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12. 245–274.

Yumoto, Yoko. 2005. *Fukugō dōshi, hasei dōshi no imi to tōgo: Mojūru keitairon kara mita nichi eigo no dōshikeisei* [Semantics and syntax of compound verbs and derived verbs: Verb formation of Japanese and English from a viewpoint of modular morphology]. Tokyo: Hituzi Syobo.

Taro Kageyama

7 Noun-compounding and noun-incorporation

1 Introduction

This chapter surveys the major patterns of compound predicates consisting of a noun as the first element, followed by a predicate (verb [V], verbal noun [VN], adjective [A], or adjectival noun [AN]) as the head, as exemplified by N-V *me-baeru* [bud-grow] 'to bud, sprout', N-VN *ronbun-sippitu* [article-write] 'write a paper', N-A *na-dakai* [name-high] 'be famous', and N-AN *Nihon-koyuu* [Japan-proper] 'proper to Japan'. Although the verbal compounds of this type might look like English "synthetic compounds" of the type *book-reading*, they are fundamentally different. While the English synthetic compounds, basically nouns or adjectives, rarely function as tensed verbs except for idiosyncratic cases of backformation (e.g. *window-shop*), the Japanese compounds under discussion function as predicates in finite or non-finite clauses. Japanese in fact has English-type synthetic compound nouns like *kane-mooke* 'moneymaking', where the noun 'money' is compounded with the deverbal noun *mooke* 'earning' (Chapter 9 [Yumoto, this volume]). In terms of their predicative function, the Japanese compound predicates might be better compared with noun incorporation phenomena in polysynthetic languages. This chapter will clarify how similar the Japanese compound predicates are to the polysynthetic noun incorporation and to the English-type synthetic compounds, as well as how different they are.

Despite the frequent use of N-V compound verbs and their kin in both written and spoken Japanese, there is scanty literature that is exclusively devoted to them – a curious lacuna that contrasts markedly with the abundance of literature dealing with Verb-Verb compound verbs (Chapter 8 [Kageyama, this volume]). This, however, by no means implies that N-V compound verbs are uninteresting. On the contrary, a close inspection of data reveals that Japanese in fact has a wide spectrum of N-V and similar compound predicates which have potential to shed new light on the general theory of noun incorporation and synthetic compounding as complex predicate formations. For the sake of brevity, I will refer to the group of Noun-Predicate compounds dealt with in this chapter as a Noun Incorporation (NI) family. It will be found that the NI family in Japanese is a constellation of NI-like compound predicates comprising a rich variety of lexical and post-syntactic compounds. Representative types are illustrated in (1).

(1) a. N + tensed V: *me-baeru* [bud-grow] 'to bud, sprout', *tema-doru* [time-take] 'to take a lot of time' (Section 2.1)

   b. N + tensed A: *ne-bukai* [root-deep] 'to be deep-rooted', *hara-guroi* [heart-black] 'to be black-hearted' (Section 2.2)

c. N + VN (so-called "post-syntactic compounds"): *Sensei ga [ronbun : sippitu] no sai* [article : write] 'when the teacher wrote a paper' (Section 3.1)

d. N + AN: *[Nihongo : koyuu] no tokutyoo* [Japanese-proper features] 'features proper to Japanese' (Section 3.2)

e. N + tenseless V: *[miti-yuku] hito* [street-go people] 'passers-by', *asemizu-tarasi-te* [sweat-drip-GER] 'working very hard' (Section 4)

f. Agent N + VN: *[Spielberg-kantoku : seisaku] no eiga* [Spielberg-director : make GEN movie] 'a movie Spielberg directed' (Section 5)

As a preliminary to detailed inspection in the ensuing sections, it is helpful to review here two archetypal theories of noun incorporation and synthetic compound formation that are pertinent to the analysis of the Japanese NI family. One is the argument-linking principle of Lieber (1983), which was intended to capture the combinatory restrictions on English synthetic compounds that were first generalized as the First Sister Principle by Roeper and Siegel (1978). In (2), I cite only the relevant portion of Lieber's principle.

(2) Lieber's (1983) argument linking:
Synthetic compounds in *-ing* are possible nouns and adjectives if the first stem is interpretable as the internal argument of the verb if it has one, or otherwise as a semantic argument if the verb lacks an obligatory internal argument.

According to (2), *ocean cruising*, for example, is interpreted as an object-verb relation ('cruise the ocean') because the verb *cruise* in this use is transitive, whereas *night cruising* is construed as an adverb-verb relation ('cruise at night') because the verb *cruise* is intransitive in this usage.

Lieber's (1983) analysis of English synthetic compounds is contrasted with Baker's (1988, 1996, 2009) purely syntactic characterization of noun incorporation in polysynthetic languages, which produces the composition of a verb or predicate with a noun in which the whole complex functions as a verb (cf. Sapir 1911; Gerdts 1998). Observe the Mohawk examples in (3), taken from Baker (1996: 12).

(3) a. *Wa'-k-hninu-'*   *ne*   *ka-nakt-a'.*
     FACT-1sS-buy-PUNC   NE   Ns-bed-NSF
     'I bought the/a bed.'

b. *Wa'-ke-nakt-a-hninu-'.*
     FACT-1sS-bed-Ø-buy-PUNC
     'I bought the/a bed.'

As opposed to the regular sentential structure of (3a), where the noun 'bed' is realized as an object phrase, the sentence in (3b) has the same noun stem *nakt* embedded within the verb complex. This latter example has an incorporation structure.

As to the combinatory possibilities of a verb and an incorporated noun, there is a solid universal tendency that the incorporated noun corresponds to the internal argument of the verb (i.e. the direct object of a transitive verb or the subject of an unaccusative verb), but never to the external argument (i.e. the subjects of transitive and unergative verbs), as reported by Mithun (1984), Baker (1988), and many other researchers. In the framework of generative grammar, Baker (1988, 1996) attributes the observed restrictions to the structural relations of syntactic structure, whereby only the argument directly governed by the verb, namely, the internal argument, is qualified for incorporation. Given that only the internal argument is subject to syntactic incorporation, to the exclusion of adjuncts as well as external arguments, it follows that the English-type synthetic compounds discussed by Lieber that permit the composition with adjuncts or semantic arguments as well as internal arguments can be regarded as lexical formations. In fact, the noun incorporations in some polysynthetic languages that permit adjunct incorporation are regarded as lexical compounding (Rosen 1989).

This chapter will provide ample evidence demonstrating that the Japanese NI family comprises both English-type synthetic compounds that are produced in the lexicon (Section 2) and polysynthetic incorporation-like compounding that takes place on syntactic structure (Section 3). The novel claim of this chapter is that these two classes are not independent of each other but are arranged on a cline defined by degrees of "non-finiteness", with non-finite or tenseless compounds being assessed as more productive than finite or tensed compounds (Sections 4 and 5).

2 Compounds of the form "N + Finite Predicate"

This section surveys the characteristics of "finite" or "tensed" compound predicates, namely those compound predicates that directly inflect for tense, negation, and other functional categories as in the present tense *tema-dor-u* [time-take-PRS] 'to take a lot of time', the past tense *tema-dot-ta* [time-take-PST] 'took a lot of time', and the present negative form *tema-dor-anai* [time-take-NEG] 'not to take a lot of time'. This kind of compound predicate, comprising N-V compound verbs and N-A compound adjectives, both from the native stratum, appears to have existed since Old Japanese (Sakakura 1966; Kinuhata 2010), although individual lexical items involved are not exactly the same at all the historical stages. In view of the paucity of previous literature providing an exhaustive list of such compounds, this section pays special attention to presenting as many examples as possible.

2.1 N-V compound verbs

We start out by clarifying the distinction between finite (or tensed) compounds and non-finite (i.e. nominal) compounds. As mentioned in Chapter 1 (Kageyama and Saito,

this volume), verb infinitives serve two functions as (i) the base for verb-selecting suffixes and (ii) the nominalized form of a native verb. The former retains the original verbal function, whereas the latter is considered a result of Verb-to-Noun conversion. The infinitive form of the verb *tur-* 'to catch', for example, is *turi*, which, as a verb, may take an object as in *sakana o turi ni* [fish ACC catch DAT] 'in order to catch fish' or be combined with the suffix *-kata* 'way' as in *turi-kata* 'the way one fishes'; as a noun, it denotes the activity of angling as in *turi ga suki da* [fishing NOM fond COP] '(I'm) fond of fishing'. This functional ambiguity of verb infinitives is manifested in compound words as well.

Compare the N-V compound verb *tabi-dat(-u)* [journey-set.out(-PRS)] 'set out on a journey' with the deverbal compound noun *sakana-turi* [fish-catching] 'fishing'. The former directly inflects for tense and negation, as in *tabi-dat-ta* [journey-set.out-PST] and *tabi-dat-anai* [journey-set.out-NEG], while the latter calls for the verb *suru* 'do' to represent tense and negation, as in *sakana-turi o si-ta* [fish-catching ACC do-PST] and *sakana-turi o si-nai* [fish-catching ACC do-NEG]. Direct attachment of tense and negation to compound nouns yields ungrammatical forms like *★sakana-tut-ta* [fish-catching-PST] and *★sakana-tur-anai* [fish-catching-NEG]. These two types of compound words also exhibit different behavior when combined with suffixes like *-kata* 'way' and *-sooda* 'look like', which select the verb category as their base (Kageyama 1982). Compound verbs are adjoined to these suffixes directly as in (4a), but compound nouns must be propped with the verb *si*, the infinitive of *suru* 'do', as in (4b).

(4)   a.   Compound verb
          tabi-dati-kata [journey-set.out-way] 'the way one sets out on a journey'
          tabi-dati-sooda [journey-set.out-look.like] 'it appears (he) is going to set out on a journey'

      b.   Deverbal compound noun
          sakana-turi no si-kata [fish-catching GEN do-way] 'the way one fishes'
          sakana-turi o si-sooda [fish-catching ACC do-look.like] 'it appears (he) is going to fish'

These observations indicate that deverbal compound nouns, like the English synthetic compounds with the *-ing* ending, are formed by combining a noun with a deverbal noun into a compound noun of the form [N [N *sakana*] – [N [V *turi*]]], and not by nominalizing a (non-existent) N-V compound verb as in ★[N [V [N *sakana*] – [V *turi*]]]. The latter structure is available only when the infinitive of a whole N-V compound verb is nominalized, as in *tabi-dati* 'departure for a journey' from *tabi-datu* [journey-set.out] and *kosi-kake* 'chair' from *kosi-kakeru* [hip-hang].

In Japanese, where case markers, especially, nominative *ga*, accusative *o*, and dative *ni*, freely undergo ellipsis in spoken language, it can be a subtle task to determine whether a given sequence of words/morphemes constitutes a word. The word-hood of the N-V compound verbs under discussion, however, can be verified rather straightforwardly by several reliable diagnoses. As Kageyama (1980, 1982) observes, the N-V complexes satisfy criteria for morphological wordhood, such as the absence of case particles, compound accent, the occasional occurrence of *rendaku* or sequential voicing, allomorphy in nouns, and subject honorification, as shown in Table 1. These characteristics are never found with the syntactic phrases that semantically correspond to the compounds.

Table 1: Word status of N-V compound verbs

	N-V compound verbs	corresponding syntactic phrases
1. Absence of case particles	*iro aseru* [color-fade] 'Colors fade' *genki zukeru* [courage-give] 'encourage'	*iro ga aseru* [color NOM fade] 'Colors fade.' *genki o tukeru* [courage ACC give] 'give courage'
2. Compound accent (Capital letters indicate high pitch.)	*iRO ASEru* [color-fade] *geNKI ZUKEru* [courage-give] 'encourage'	*iRO ga aSEru* [color NOM fade] *GEnki o tuKEru* [courage ACC give]
3. *rendaku* or voicing of the initial consonant of the second member	*tumasaki datu* [tiptoe-stand] 'stand on tiptoe' < *tatu* 'stand' *me buku* [bud-shoot.forth] 'bud, sprout' < *huku* 'shoot forth'	*tumasaki de tatu* [tiptoe INST stand] *me ga huku* [bud NOM shoot.forth]
4. Allomorphy in nouns	*ta oru* [hand-break] 'snip off by hand' < *te* 'hand' *tuma biku* [nail-pluck] 'pick (a guitar)' < *tume* 'nail'	*te de oru* [hand INST break] *tume de hiku* [nail INST pluck]
5. subject honorification (*o*-V *ni naru*)	*Sensei wa o tabi dati ni nat-ta.* [teacher TOP HON-journey-set. out DAT become-PST] 'The teacher set out on a journey.'	**Sensei wa o tabi ni tati ni nat-ta.* [teacher TOP HON-journey DAT set.out DAT become-PST]

In addition, contraction reinforces the phonological unity of some compounds. For example, *se-ou* [shoulder-carry] 'carry on one's back' is contracted to *syou* in casual speech.

Evidence for the lexicalized status of the N-V compound verbs is not hard to come by. First, certain N elements, like the noun allomorphs with vowel change in Table 1, have changed to bound morphemes due to loss of their original meanings, as shown in (5).

(5) **kaima**-*miru* [fence.slit-see] 'peep through, catch a glimpse of', **saka**-*noboru* [reverse-go.up] 'trace back', **una**-*dareru* [head-droop] 'droop one's head', **me**-*kurumeku* [eye-swirl] 'be dazzling', **ta**-*mukeru* [hand-direct] 'offer to the deceased/god', **me**-*zasu* [eye-point] 'aim (at)', **me**-*gakeru* [eye-set] 'aim (at)', **to**-*daeru* [track-end] 'be cut off', **omo**-*muku* [face-turn] 'go to', **ukiasi**-*datu* [floating.leg-stand] 'become upset and ready to run away', **ha**-*mukau* [blade-turn.toward] 'rebel against'

The loss of transparency in morphological composition is sometimes signaled by orthography, as witness some fossilized verbs that are written with an arbitrary *kanji* character that lacks etymological motivation. Native speakers will not easily recognize the etymology of the examples in (6) as N-V compounds.

(6) 育む *ha-gukumu* [feather-wrap] 'to bring up, nurture (as a bird broods over its chicks)', 頷く *una-zuku* [head-thrust] 'nod', 躓く *tuma-zuku* [toe.nail-thrust] 'stumble, trip over', 滴る *sita-taru* [down-drop] 'drip', 跪く *hizama-zuku* [knee.head-attach] 'kneel down', 偏る *kata-yoru* [one.side-lean] 'be biased', 助ける *ta-sukeru* [hand-save] 'help', 背く *so-muku* [back-turn] 'betray', 耕す *ta-gayasu* [field-turn] 'cultivate'

Second, compounds tend to acquire specialized meanings that are not available to their phrasal counterparts. For example, *himo-toku* [string-untie] lit. 'untie the string that binds a roll' is now commonly used to mean 'to read (a book)' and by further extension 'examine in detail', *sao-sasu* [pole-push] lit. 'to push a long pole against the bottom of the river' can be used to mean 'to swim with the tide', *tabi-datu* [journey-set.out] 'start a new life', *iro-aseru* [color-fade] 'lose the original freshness', *iki-maku* [breath-whirl] 'be infuriated', and *kokoro-gakeru* [mind-hang] 'endeavor to'. Additionally, it often happens that the meaning of a compound verb is limited to a metaphorical meaning while its phrasal counterpart is ambiguous between a literal and an extended meaning, as in the compound *hone-oru* [bone-break] expressing only the extended meaning of 'to take pains, endeavor' in contrast to the verb phrase *hone o oru* [bone ACC break], which may mean either 'to fracture a bone' or 'to take pains'.

Third, and most importantly, the combinations of individual nouns and individual verbs are lexically predetermined, and no principled explanation for lexical gaps seems available. For example, **kata.asi-datu* [one.foot-stand] 'to stand on one leg' is a lexical gap as opposed to the existing *tumasaki-datu* [tiptoe-stand] 'to stand on tiptoe'. Note that *kata.asi-dati* [one.foot-standing], *tumasaki-dati* [tiptoe-standing], and other deverbal compound nouns of the form 'N-*dati*' [N-standing] are all possible – a fact that demonstrates that the formation of N-V compound verbs is severely limited whereas that of deverbal compound nouns is fully productive. The

limited productivity of N-V compound verbs, however, does not preclude the possibility of the emergence of novel N-V compound verbs, as illustrated by compound verbs of the form 'N-*zukeru*' [N-add] such as *genki-zukeru* [vigor-add] 'invigorate', *kankei-zukeru* [relation-add] 'relate', and *yuuki-zukeru* [courage-add] 'encourage' – a pattern that was born only in the early 20th century, according to the *Nihon kokugo daijiten* (2000–2002).

Having observed the lexical and morphological idiosyncrasies of N-V compound verbs, we are in a position to probe the internal semantic and grammatical relations between the V and N elements. Except for those N-V compounds that are fossilized as single verbs, as in (6), the N-V compound verbs are analytic with respect to their internal relations in the sense that they can be roughly paraphrased by using clausal structure with appropriate case particles on the nouns, as in *me-zameru* [eye-awake] 'wake up' being paraphrased as *me ga sameru* [eye NOM awake] in an subject-unaccusative verb relation, *tema-doru* [time-take] 'take lots of time' paraphrasable as *tema o toru* [time ACC take] in an object-transitive verb relation, and *tumasaki-datu* [tiptoe-stand] 'stand on tiptoe' paraphrasable as *tumasaki de tatu* [tiptoe INST stand] in an adjunct-verb relation. In my database comprising sixty-five N-V compound verbs, their semantic and grammatical relations are roughly distributed as shown in Table 2.

Table 2: Proportion of semantic and grammatical relations in N-V compound verbs

	type count	percentage
Unaccusative subjects	15	23%
Transitive objects	24	37%
Dative complements	4	6%
Adjuncts	22	34%
Transitive subjects	0	0%
Dative objects	0	0%
Total	65	100%

As seen from Table 2, the total number of N-V compound verbs that I have collected for this paper is not so large, and it is difficult to form new ones. This means that N-V compound formation is not a productive word formation process in Japanese.

The predominant relation is the internal argument relation, where the N element functions as the direct object if the V element is transitive or as the subject if it is unaccusative. Examples of unaccusative subjects are shown in (7), and examples of direct objects in (8).

(7) Subjects of unaccusative verbs (paraphrasable as 'N *ga* [NOM] V')
me-baeru [bud-grow] 'bud, sprout', *me-buku* [bud-shoot] 'bud, sprout',
me-zameru [eye-awake] 'wake up', *iro-zuku* [color-change] '(tree leaves) change
colors', *kizu-tuku* [injury-be.attached] 'be injured', *ki-zuku* [attention-get.to]
'notice', *nami-datu* [wave-rise] 'waves rise, billow', *nami-utu* [wave-hit] 'surge',
awa-datu [bubble-rise] 'bubble, foam', *uzu-maku* [whirlpool-whirl] 'whirl',
iro-aseru [color-fade] 'colors fade', *iki-taeru* [breath-stop] 'die', *hara-datu*
[stomach-stand] 'feel angry', *genki-zuku* [vigor-come.out] 'be invigorated',
hana-hiraku [flower-open] 'come into blossom'

(8) Objects of transitive verbs (paraphrasable as 'N *o* [ACC] V')
syuukan-zukeru [habit-assume] 'acquire the habit', *mi-gamaeru* [posture-take]
'take a posture', *tema-doru* [time-take] 'take lots of time', *na-zukeru* [name-attach]
'name', *sao-sasu* [pole-put.in] 'swim with the tide', *muti-utu* [whip-apply]
'whip', *ada-nasu* [revenge-take] 'take revenge', *ki-zukau* [care-take] 'be con-
cerned about', *tema-doru* [time-take] 'take lots of time', *tosi-toru* [age-take]
'get old', *kosi-kakeru* [hip-put.down] 'sit', *hone-oru* [bone-break] 'take pains',
zin-doru [camp-set.up] 'set up a camp', *mei-utu* [name-give] 'name, label', *yuuki-*
zukeru [courage-attach] 'encourage', *genki-zukeru* [vigor-attach] 'invigorate',
kankei-zukeru [relation-attach] 'relate', *iti-zukeru* [location-attach] 'locate',
tokutyoo-zukeru [character-attach] 'characterize', *sei-dasu* [efforts-make], 'make
efforts', *te-gakeru* [hand-put] 'put one's hand to', *kokoro-ubawareru* [heart-
deprive.PASS] 'be captivated', *kokoro-hikareru* [heart-attract.PASS]
'be fascinated'

The last two examples in (8), *kokoro-ubawareru* 'be captivated' and *kokoro-hikareru*
'be fascinated', pronounced with a slight pause between the two members, are rare
examples consisting of a passive verb and its direct object. These are also special in
that their active counterparts **kokoro-ubau* [heart-deprive] and **kokoro-hiku* [heart-
attract] are unavailable in finite forms and are restricted to adnominal usage (see
Section 4).

A small number of compound verbs involving dative complements, which can
be assimilated to the group of internal argument relations because the dative com-
plements show up directly before a head verb, are shown in (9), and diverse adjunct
relations are shown in (10).

(9) Dative or locative relations (paraphrasable as 'N *ni* [DAT] V')
tabi-datu [journey-set.out] 'set out on a journey', *se-ou* [back-carry] 'carry on
one's back', *yaku-datu* [purpose-serve] 'be useful', *mi-gomoru* [body-bear]
'bear a baby', *su-gomoru* [nest-be.confined] 'stay in the nest'

(10) Adjunct relations (instrumental, direction, path, and other adverbial relations)
tumasaki-datu [tiptoe-stand] 'stand on tiptoe', *yume-miru* [dream-see] 'dream of',
kaima-miru [slit-see] 'catch a glimpse of', *kaima-mieru* [slit-be.seen] 'be vaguely

seen', *te-maneku* [hand-beckon] 'beckon', *me-datu* [eye-stand] 'be conspicuous', *uwa-mawaru* [above-turn] 'exceed', *sita-mawaru* [below-turn] 'fall below', *yoko-giru* [sideways-cut] 'cross', *yoko-muku* [sideways-turn] 'turn sideways', *saka-noboru* [backward-climb] 'trace back', *sita-taru* [down-drop] 'drip', *saki-datu* [ahead-go] 'go ahead of, precede', *tika-zuku* [near-arrive] 'draw near', *too-zakaru* [far-leave] 'go away', *yubi-sasu* [finger-point] 'point at', *kusi-kezuru* [comb-scratch] 'comb', *te-watasu* [hand-give] 'hand', *tuma-biku* [nail-pluck] 'pick (a guitar)', *hara-bau* [stomach-crawl] 'lie on one's stomach', *te-banasu* [hand-let.go] 'part with, dispose of', *ta-basamu* [hand-hold] 'hold under one's arm'

As opposed to internal arguments and adjuncts, no unequivocal example is attested that involves incorporation of external arguments, i.e. subjects of transitive or unergative verbs. A possible counterexample is *musi-bamu* [bug-bite], in which the N *musi* 'bug' looks like the subject of the transitive verb 'bite'. The noun 'bug', however, should be interpreted as a kind of manner adverb meaning 'like a bug, as if by a bug', whereby the whole compound conveys a metaphorical meaning of 'to erode' or 'to affect adversely' as in *Gan ga kanozyo no karada o musi-ban-da.* [cancer NOM she GEN body ACC bug-bite-PST] 'Cancer eroded her body'. Another apparent counterexample is *kami-gakaru* [god-obsess] 'become obsessed, be possessed by a supernatural spirit', which is normally used as an adnominal modifier instead of a tensed verb (cf. Section 4). Although the conformity with the internal argument restriction might suggest a syntactic derivation by incorporation in the sense of Baker (1988, 1996), a variety of semantic relations observed in (10) strongly favor lexical derivation of N-V compound verbs, on a par with English synthetic compounds as regulated by Liber's linking principle in (2).

In previous literature, attempts were made to derive N-V compound verbs from syntactic structure. One, Kageyama (1982), brings two facts to bear on the lexical-syntactic distinction. First, the compounding does not affect the grammatical relations of the "unincorporated" arguments, as shown by (11), and second, the nouns, once incorporated into verbs, cannot show up in duplicate in the sentence (so-called "doubling"), as in (12).

(11) a. *Syatyoo   wa   isu   ni   kosi   o   kake-ta.*
          president   TOP   chair DAT   hip   ACC   put.down-PST
          'The president sat on the chair.'

      b. *Syatyoo   wa   isu   {ni/*o}   [ᵥ kosi-kake]-ta.*
          president   TOP   chair {DAT/ACC}   [hip-put.down]-PST

(12) **Syatyoo   wa   isu   ni   ookii   **kosi**   o   [**kosi**-kake]-ta*
       president   TOP   chair DAT   large   hip   ACC   [hip-put.down]-PST
       lit. 'The president sat his large hips on the chair.'

The retention of the original argument relation in (11) and the ban on doubling in (12) appear to be parallel to the conditions on syntactic noun incorporation in polysynthetic languages (Baker 1988; Rosen 1989).

Upon closer inspection, however, examples are found in which these conditions are not met. (13) illustrates examples of doubling, and (14) examples of argument relation changing (Kageyama 1980).

(13) a. *na   o   [na-noru]*
       name   ACC   [name-tell]   'tell one's name'

   b. *tuma   o   [me-toru]*
       wife   ACC   [wife-take]   'take a wife'
       tuma: contemporary Japanese 'wife', *me*: Old Japanese 'woman, wife'

(14) a. *Mizuumi   ni   nami   ga   tatu. / Mizuumi {ga/*ni}   [nami-datu].*
       lake   DAT   wave   NOM   rise   lake   {NOM/*DAT}   [wave-rise]
       'Waves rise on the lake.'

   b. *Iwa   ni   koke   ga   musu. / Iwa   {ga/*ni}   [koke-musu].*
       rock   DAT   moss   NOM   grow   rock   {NOM/*DAT}   [moss-grow]
       'Moss grows on the rock.'

These examples take us back to the lexical derivation. The plausibility of lexical analysis is enhanced by the inability of an incorporated noun to participate in sentence-level anaphora, as shown in (15).

(15) *Otto   wa   tuma   ni   purezento   o   [te$_i$-watasi]-ta.*
      husband   TOP   wife   DAT   gift   ACC   [hand-give]-PST.

   *Sore$_i$ wa   yogorete   ita.
    It   TOP   dirty   was.
    'The husband handed his wife a gift, but it was dirty.'

In (15), the pronoun *sore* 'it' may refer to the object 'gift', but never to the noun 'hand' contained in the compound verb.

Table 3 summarizes the grammatical properties of incorporated nouns in Japanese N-V compounds as contrasted with those in polysynthetic noun incorporation.

Table 3: Characteristics of incorporated/compounded nouns

	Japanese N-V verbs	**polysynthetic NI**
definiteness	indefinite	indefinite
referentiality	non-referential	referential
argument relations	internal argument, adjunct	internal argument
doubling	allowed in some	no

Indefiniteness will fall out from the morphological word status of a whole compound, whereas referentiality and argument relations are directly associated with syntactic structures. While the incorporation in those polysynthetic languages in which the incorporated nouns are referential and are structurally limited to the internal argument of the host verb is legitimately considered syntactic, the N-V compounding of the native Japanese stratum is best considered lexical.

The validity of the lexical treatment is confirmed crucially by the total prohibition on the deletion of nominal elements in coreferential environments like those in (16).

(16) a. A: *Ki   wa   iro$_i$ -zuki   masi-ta ka?*   B: **Hai, Ø$_i$ tuki   masi-ta.*
        tree TOP color$_i$ -change POL-PST Q   Yes, Ø$_i$ change POL-PST
        'Have the tree leaves changed colors?'   'Yes, they have.'

    b. A: *Syukudai   wa   tema$_i$-dori masi-ta ka?*   B: **Hai, Ø$_i$ tori masi-ta.*
        assignment TOP time-take POL-PST Q   Yes, Ø$_i$ take POL-PST
        'Did the assignment take lots of time?'   'Yes, it did.'

In (16), B's utterances do not allow the N elements of the compounds to be deleted under identity with the same nouns in A's utterances. This is because syntactic deletion cannot affect part of lexical words ("lexical integrity principle"). The total ungrammaticality of B's utterances in (16) should be contrasted with the total grammaticality of (17) involving verb complexes consisting of VNs and *suru* 'do' (Kageyama 1982).

(17) *Gakkai   de   Amerika-zin   wa   yoku   hatugen$_i$-su-ru   ga,*
    conference LOC   American-people   TOP often   remark-do-PRS but

    Nihon-zin   wa   amari Ø$_i$ si-nai.
    Japanese-people TOP rarely Ø   do-NEG
    'At conferences, Americans always speak out, but Japanese people seldom do so.'

The sentence in (17), sharing the same structure as (16a, b), is perfectly grammatical despite the fact that it has no N element on the verb of the second clause. The verb in the second clause, *si-nai* 'don't', is easily understood as an omission of *hatugen* 'remark' from the compound verb *hatugen-suru* [remark-do] 'speak out' that shows up in the first clause. This is because, according to Kageyama (1982), the VN and *suru* are compounded at a certain level of syntax, thereby evading the lexical integrity effect.

It is also instructive to compare these native N-V compound verbs with two-morpheme compound VNs of the Sino-Japanese stratum. Since Sino-Japanese VNs form a tight morphological unit consisting of two bound roots (one representing a

noun concept and the other a verb concept), it would appear most plausible to derive them by lexical compounding. Nevertheless, there are some facts that favor syntactic derivation. Consider the behavior of *syuk-ka* [break.out-fire] 'fire breaks out', consisting of an unaccusative verbal ('break out') and its subject noun ('fire'), with respect to grammatical-relation changing and doubling (cf. Kageyama 1980, 1993).

(18) a. *Daremo   i-nai   heya   {kara/?*ga}   [syuk-ka]-si-ta.*
          anyone   be-NEG   room   {ABL/*NOM}   [break.out-fire]-do-PST
          'Fire broke out from a room where there was supposed to be nobody.'

    b. **Daidokoro   kara   hi   ga   [syuk-ka]-si-ta.*
          kitchen   from   fire   NOM   [break.out-fire]-do-PST
          'Fire broke out from the kitchen.'
          hi: native noun 'fire', *ka*: S-J noun 'fire'

As shown by (18a), *syuk-ka* does not affect the grammatical relation of the sentence, leaving the source phrase ('a room') with the ablative marking. This example itself is interesting because it displays no nominative phrase in violation of the usual assumption that a sentence must contain at least one nominative phrase (Shibatani 1978). Furthermore in (18b), *syuk-ka* does not allow doubling of the noun 'fire' as the subject. Even stronger evidence suggestive of syntactic derivation is found in (19) below, where the N element in the compound serves as the antecedent of a subsequent anaphoric expression (Kageyama 1993).

(19) *Daidokoro   kara   [syuk-ka$_i$]-si-te,   Ø$_i$ tatimati   ie-zyuu   ni*
       kitchen   ABL   [break.out-fire$_i$]-do-GER *pro$_i$* instantly house-all.over DAT

    moe-hirogat-ta.
       burn-spread-PST
       'Fire broke out from the kitchen and spread all over the house in a flash.'

In (19), the verb *moe-hirogaru* 'to spread burning' in the second conjunct superficially lacks its subject. Semantic consideration indicates, however, that its empty subject should refer to the bound root *-ka* 'fire' contained inside the S-J compound *syuk-ka*. The same kinds of arguments can be made for the VN *raku-rai* [fall-thunderbolt] 'a thunderbolt strikes', as in *Sakuya kooen {ni/*ga} raku-rai-si-ta* [last.night park {LOC/*NOM} fall-thunderbolt-do-PST] 'A thunderbolt struck the park last night'.

    The above two examples, *syuk-ka* 'fire breaks out' and *raku-rai* 'thunderbolt strikes', have a particularly transparent internal structure composed of an unaccusative verb and a subject noun, and this semantic and morphological transparency undoubtedly contributes to activating the function of the word-internal nouns, *ka* 'fire' and *rai* 'thunderbolt', as syntactic subjects. If the syntactic visibility attested

in these phenomena is taken as evidence for syntactic formation, it follows that at least these two S-J verbal nouns are formed in the syntax, although the technical implementation is yet to be explored. One such attempt is made by Watanabe (1999), where a syntactic head-movement analysis in line with Hale and Keyser's (2002) lexical syntactic structure with layered VP shells is suggested uniformly for both the native N-V compound verbs and the Sino-Japanese VNs without any syntactic evidence like the anaphoric behavior of the incorporated Ns. As we observed above, a unitary syntactic treatment of native N-V compound verbs and Sino-Japanese VNs fails to capture their different behavior in anaphora and other phenomena. This comment applies with equal force to the theory of Distributed Morphology, which aims to subsume all word formation processes in syntax.

To summarize, the N-V compound verbs of the native stratum should be viewed as a lexical derivation, whereas (some) two-character Sino-Japanese VNs leave room for syntactic analysis – a suggestion that sounds odd in light of the bound nature of their component elements. In Section 3, we will observe another idiosyncratic behavior of Sino-Japanese VNs under the heading of "post-syntactic compounding".

2.2 N-A compound adjectives

The Japanese adjectives, all in the native stratum, inflect for tense by themselves, so that they perform the same syntactic function as inflected verbs, taking subjects and complements of their own. From this, it would be expected that N-A compound adjectives display much the same behavior as native N-V compound verbs. Indeed, the facts are as expected.

As is the case with N-V compound verbs, the absence of case particles on the N elements is an initial indicator of the wordhood of N-A compounds. Compare the compound *ne-bukai* [root-deep] 'deep-rooted' with the clause *ne ga hukai* [root NOM deep] 'The root is deep' marking the subject *ne* 'root' with the nominative case *ga*. The compound status of *ne-bukai* is confirmed by the voicing of the initial consonant of *hukai* 'deep' to *bukai* – [h] alternates with [b] because it was supposedly [p] in Old or Proto Japanese. Curiously, no N-A compound adjectives are found that include noun allomorphy, like the alternation of *te* 'hand' and *ta-* in N-V compound verbs (e.g. *ta-oru* 'snap by hand'). This gap, however, is partially complemented by prefix-like nouns that have lost their literal meaning and express degrees of the properties denoted by adjectives, as exemplified by (20). There are also lexicalized compound adjectives carrying specialized meanings where the semantic or grammatical relations between the N and A elements are not completely clear, as in (21).

(20)   *mono-ganasii* [thing-sad] 'somehow sad', *ura-sabisii* [heart-lonely] 'somehow lonely', *ki-hazukasii* [heart-ashamed] 'somehow ashamed', *kokoro-hazukasii* [heart-ashamed] 'somehow ashamed', *kuti-osii* [mouth-regretful] 'very regretful'

(21)　*te-kibisii* [hand-severe] 'unsparing, bitter', *kokoro-bosoi* [heart-thin] 'uneasy',
　　　kokoro-zuyoi [heart-strong] 'reassuring', *me-zamasii* [eye-awakening] 'remark-
　　　able', *me-atarasii* [eye-new] 'novel', *keisan-dakai* [calculation-high] 'calculat-
　　　ing, mercenary', *te-arai* [hand-rough] 'rough in behavior', *ne-zuyoi* [root-strong]
　　　'firm', *kuti-sabisii* [mouth-insufficient] 'feel like eating', *iki-gurusii* [breath-hard]
　　　'suffocating', *reigi-tadasii* [manners-right] 'courteous, well-mannered',
　　　hodo-tooi [distance-far] 'have a long way to go', *omo-hayui* [face-shiny]
　　　'be embarrassed'

The examples in (20) involve psychological adjectives, which are generally assumed
to take human experiencers, i.e. external arguments, as subjects. Because of this,
Lieber's argument linking principle correctly precludes the interpretation of the N
elements as subjects, leaving them only with prefix-like interpretations. Likewise in
(21), although the N elements appear to retain their literal meanings somehow, the
whole compounds have developed idiosyncratic meanings. All of the N-A adjectives
in (20) and (21) will be characterized as stage-level predicates that denote a transi-
tory and unstable state of the subject NP they are predicated of.

　　Aside from these more or less opaque cases, there are analytical N-A compound
adjectives whose internal relation is limited to internal arguments (cf. Namiki 1988;
Yumoto 1990; Kageyama 1993). These compounds are thus paraphrased with the
construction 'N *ga* [NOM] A', as in *ne-bukai* [root-deep] paraphrased as *ne ga hukai*
'the root is deep'. As seen from the classified examples in (22), the productivity of
this type of compound is largely dependent upon individual head adjectives.

(22)　Subject-adjective relations (N *ga* [NOM] A.)
　　　a.　N-*hukai/bukai* [N-deep]
　　　　　ne-bukai [root-deep] 'deep-rooted', *oku-hukai/bukai* [inner.part-deep]
　　　　　'profound', *ke-bukai* [hair-thick] 'thick-haired, hairy', *tumi-bukai* [sin-deep]
　　　　　'sinful', *sinzin-bukai* [piety-deep] 'pious', *insyoo-bukai* [impression-deep]
　　　　　'impressive', *enryo-bukai* [modesty-deep] 'modest, reserved', *tutusimi-bukai*
　　　　　[restraint-deep] 'discrete, modest', *syuunen-bukai* [attachment-deep]
　　　　　'persistent', *sitto-bukai* [jealousy-deep] 'jealous', *siryo-bukai* [discretion-
　　　　　deep] 'discrete'

　　　b.　N-*takai/dakai* [N-high]
　　　　　kakutyoo-takai [tone-high] 'high-toned, stylish', *na-dakai* [name-high]
　　　　　'famous', *kasa-dakai* [volume-high] 'bulky', *kaori-takai* [fragrance-high]
　　　　　'fragrant', *hokori-takai* [pride-high] 'proud',

　　　c.　N-*yoi/ii* [N-good]
　　　　　kakkoo-yoi/ii [style-good] 'stylish, cool', *kimoti-yoi/ii* [feeling-good]
　　　　　'pleasant, agreeable', *tugoo-yoi* [convenience-good] 'convenient'

d. N-*warui* [N-bad]
   izi-warui [nature-dirty] 'ill-natured', *kakkoo-warui* [style-bad] 'unstylish, uncool', *tugoo-warui* [convenience-bad] 'inconvenient', *kimoti-warui* [feeling-bad] 'unpleasant', *kimi-warui* [feeling-bad] 'weird, creepy'

e. N-*zuyoi* [N-strong]
   gaman-zuyoi [endurance-strong] 'patient', *sinboo-zuyoi* [patience-strong] 'patient', *nebari-zuyoi* [tenacity-strong] 'tenacious'

f. N-*nai* [N-null]
   azike-nai [taste-null] 'tasteless', *darasi-nai* [moderation-null] 'slovenly', *sikata-nai* [way.out-null] 'inevitable'

g. Other heads
   en-dooi [relation-distant] 'not closely related', *hara-guroi* [heart-black] 'black-hearted', *kime-komakai* [grain-fine] 'fine-grained, *te-bayai* [hand-quick] 'quick in movement', *nagori-osii* [parting-reluctant] 'reluctant to part/leave', *kuti-gitanai* [mouth-dirty] 'foul-mouthed, abusive', *izi-kitanai* [nature-dirty] 'greedy', *kazu-ooi* [number-numerous] 'numerous', *kokoro-yasasii* [heart-kind] 'kind-hearted'

Most compound adjectives of this type are individual-level predicates characterizing a more or less permanent and stable property of the NP they are predicated of. Exceptions are very few, as in *tugoo-yoi* [convenience-good] 'convenient' and *tugoo-warui* [convenience-bad] 'inconvenient'.

Because of their lexical idiosyncrasies, such as adjunct and prefixal interpretations as well as development of new meanings, all the N-A compounds that have been discussed so far are subject to lexical treatment. As regards the lexical-syntactic distinction, Kishimoto and Booij (2014) bring to light an interesting set of compound adjectives headed by the negative adjective *nai* 'non-existent, null'. Observe the two classes of expressions in (23) and (24).

(23) *sikata (ga) nai* [way.out (NOM) null-PRS] 'unavoidable'
      moosiwake (ga) nai [excuse (NOM) null-PRS] 'inexcusable'
      darasi (ga) nai [tidiness (NOM) null-PRS] 'untidy'
      otonage (ga) nai [grownup (NOM) null-PRS] 'childish'
      moosibun (ga) nai [complaint (NOM) null-PRS] 'perfect'
      itasikata (ga) nai [way.out (NOM) null-PRS] 'cannot help'

(24) *yurugi (ga) nai* [shake (NOM) null-PRS] 'firm, unshakable'
      atogusare (ga) nai [later.concern (NOM) null-PRS] 'without later trouble'
      nukari (ga) nai [oversight (NOM) null-PRS] 'shrewd'
      abunage (ga) nai [danger (NOM) null-PRS] 'safe, without danger'
      sonsyoku (ga) nai [inferiority (NOM) null-PRS] 'comparable'

The expressions in (23) and (24) share the syntactic construction 'N *ga* [NOM] *nai*' (lit. 'N is absent/nonexistent'). They are syntactic adjectival phrases when their N elements are marked with the nominative particle *ga*, but make up compound adjectives when the case particle drops out, as shown by the fact that the nominalizing suffix -*sa* applies only in the absence of the nominative particle: **yurugi ga na-sa* [shake NOM null-ness] vs. *yurugi-na-sa* [shake-null-ness] 'unshakableness, firmness'. The availability of the nominative case particle in paraphrase, however, does not necessarily guarantee that the compound adjectives in both (23) and (24) are formed directly from their respective case-marked syntactic constructions. Consider the possibility of deleting only the N element from the compound, which is impossible not only with the compound adjectives discussed earlier, as in (25a), but also with the compound negatives in (23), as shown by (25b).

(25) a. A: *Kare wa   ke$_i$-bukai desu ka?*   B: **Hai, Ø$_i$ hukai desu.*
            he   TOP hair-deep COP Q   Yes,   deep   COP
            'Is he hairy?'   'Yes, he is.'

    b. A: *Ano ziko   wa   sikata$_i$-nai   ne?*   B: **Un, Ø$_i$ nai   ne.*
            that accident TOP doing.way-null SFP   Yes,   null SFP
            'That accident was unavoidable."   'Yes, it was.'

Since the N elements cannot be detached from the head adjectives, the compound adjectives given in (23) are identified as holistic words in syntactic structure rather than being derived from the corresponding case-marked constructions.

By contrast, the compound adjectives corresponding to the adjectival phrases in (24) are amenable to syntactic formation – a process which Kishimoto and Booij (2014) call Quasi Noun Incorporation. Although they do not provide concrete arguments demonstrating that the compounds corresponding to (24) originate from syntactic structure, dialogs like (26) involving omission of N elements provide strong evidence that the N elements and *nai* are separate words at some point in syntactic derivation.

(26) a. A: *Atogusare$_i$-nai   ka?*   B: *Un, Ø$_i$ nai   yo.*
            later.concern-null   Q   Yes,   null SFP
            'Won't there be no trouble later?'   'No, there won't.'

    b. A: *Nukari$_i$-nai   ka?*   B: *Un, Ø$_i$ nai   yo.*
            fault-null   Q   Yes,   null SFP
            'Isn't there any oversight?'   'No, there isn't.'

In B's responses of (26a, b), *nai* 'null' takes a null subject (Ø) that refers to the N elements (*atogusare* 'trouble arising later' and *nukari* 'oversight') in the compound

adjectives used in A's questions. The same argument can be made with the polite form of *nai*, i.e. *ari-mas-en* [be-POL-NEG] as well.

Before closing this section, we draw attention to a group of N-A adjectives that call for a separate treatment from the N-A compounds discussed so far.

(27)   *kabi-kusai* [mold-smelly] 'stinking of mold', *ase-kusai* [sweat-smelly] 'stinking of sweat', *ninniku-kusai* [garlic-smelly] 'stinking of garlic', *sake-kusai* [alcohol-smelly] 'stinking of alcohol', *gasorin-kusai* [gasoline-smelly] 'stinking of gasoline', *sakana-kusai* [fish-smelly] 'stinking of fish'

Compounds of type (27), created productively to represent a particular kind of bad smell typically associated with the noun elements, do not fit into the argument-structural characterization with the subject relation. For example, *kabi-kusai* [mold-smelly] does not mean *kabi ga kusai* [mold NOM smelly] 'The mold smells', because what is smelly is not the noun entity ('mold') within the compound but the entity that is predicated of by the compound adjectives, as in *Kono heya wa kabi-kusai* [This room TOP [mold-smelly]] 'This room smells of mold'. For this sentence to be true, there need not be mold in the room; instead it is only necessary that the room has the kind of smell that is typically associated with mold. An even clearer example is *tosiyori-kusai* [elderly.man-smelly] 'be like an elderly man', which might be plausibly used to characterize even a young person. The meaning of the N-*kusai* 'smelly' compounds will thus be represented not as a subject-predicate relation but as λx.**give.off**´ (**e**, x, **a smell characteristic of** N), namely, 'giving off a smell charac-teristically associated with N'. This formula applies also to cases where -*kusai* is compounded with a verb, as in *koge-kusai* [burn-smelly] 'giving off a smell of some-thing burning'.

3 Compounds of the form "N + VN/AN"

This section gives an overview of what Shibatani and Kageyama (1988) termed "post-syntactic compounds" and their related compound predicates, which all share the structure 'N+VN (verbal noun)' or 'N+AN (adjectival noun)'. Since these compounds are fully productive with virtually no lexical restrictions, only the core properties will be explained with a handful of illustrative examples.

3.1 N-VN compound predicates

Post-syntactic compounding, discovered by Shibatani and Kageyama (1988) and ela-borated on by Kageyama and Shibatani (1989) and Kageyama (1993), characteristi-cally apply only in certain syntactic constructions, as exemplified by (28).

(28) a. Tenseless clause structure

Sakuzitu	*syusyoo*	*ga*	[*hisaiti* : *sisatu*]	*no*	*sai*
yesterday	prime.minister	NOM	[stricken.area : inspect]	GEN	time

'When the Prime Minister visited/visits the disaster-stricken area, …'

b. Noun phrase structure

Sakuzitu	*no*	*syusyoo*	*no*	[*hisaiti* : *sisatu*]	*wa*
yesterday	GEN	prime.minister	GEN	[stricken.area : inspection]	TOP

'The Prime Minister's visit to the disaster-stricken areas yesterday was …'

In (28a, b), the square brackets indicate post-syntactic compounds, with the symbol ":" showing a slight phonological break that characteristically occurs inside them. Because of the phonological break, the compound [*hisaiti* : *sisatu*] is pronounced with accents on the two components, *hiSAITI* and *siSATU*, instead of a single accent characteristic of lexical compounds. This kind of compound is presented as a unique phenomenon in the world's languages by Spencer (1991) and Fabb (1998) in their introductory treatises.

As shown by the two examples, post-syntactic compounding takes place both in tenseless adverbial clauses with regular case markers on the arguments and adjuncts expressed in the same clauses, as in (28a), or in noun phrase structures with the arguments and adjuncts marked with the genitive, as in (28b). The compounds in (28a) and (28b) are respectively analyzed as resulting from their corresponding clausal or phrasal structures in (29a) and (29b).

(29) a. | *Sakuzitu* | *syusyoo* | *ga* | *hisaiti* | *o* | *sisatu* | *no* | *sai* |
|---|---|---|---|---|---|---|---|
| yesterday | prime.minister | NOM | stricken.area | ACC | inspect | GEN | time |

'When the Prime Minister visited/visits the disaster-stricken area, …'

b. | *Sakuzitu* | *no* | *syusyoo* | *no* | *hisaiti* | *no* | *sisatu* | *wa* |
|---|---|---|---|---|---|---|---|
| yesterday | GEN | prime.minister | GEN | stricken.area | GEN | inspection | TOP |

'The Prime Minister's visit to the disaster-stricken areas yesterday was …'

The semantic interpretations of the compounds in (28a) and (28b) are precisely the same as those of the corresponding phrasal structures in (29a) and (29b).

A plausible objection here would be that the bracketed parts in (28) are mere cases of particle drop, which is extremely common in Japanese. This objection, however, does not go through. First, particle drop typically occurs in colloquial speech, whereas post-syntactic compounding is characteristically found in elevated speech and journalistic writings. Second, the genitive particle, which generally cannot be dropped out even in colloquial speech (e.g. *syusyoo *(no) kaban* [prime.minister *(GEN) bag] 'the Prime Minister's bag', is legitimately elided in post-syntactic compounds in the noun phrase structure, as in (28b).

Beside the absence of case particles, there is additional evidence indicating the word status of post-syntactic compounds. The examples in (30) show that syntactic elements like adverbs and nominal modifiers are prohibited from occurring inside the compounds (lexical integrity principle), and those in (31) demonstrate that two (or more) nouns cannot be compounded with a single head (Selkirk's (1982) "binary branching condition": word structure is generally built by combining of two elements).

(30) a. *Syusyoo   ga   [hisaiti :   _totuzen_   sisatu]   no   sai
          prime.minister NOM   [stricken.area : <u>suddenly</u> inspect] GEN   time
          'When the Prime Minister suddenly inspected/inspects the stricken area'

    b. *Syusyoo   no   [hisaiti : _totuzen no_ sisatu]   wa
          prime.minister GEN   [stricken.area : <u>sudden GEN</u> inspection] TOP

(31) a. Kaityoo   ga   yuusyoosya ni   [kin-medaru : zyuyo]   go
          president   NOM   winner   DAT [gold-medal : give]   after
          'After the president gave/gives a gold medal to the winner'

    b. *Kaityoo   ga   [yuusyoosya : kin-medaru : zyuyo]   go
          president   NOM   [winner : gold-medal : give]   after

Colloquial ellipsis of case particles is exempt from such morphological restrictions.

A conspicuous restriction of post-syntactic compounding is that it applies only to bare VNs involving no tense carrier like _suru_ 'do' or its morphological suppletions (_dekiru_ 'be able to' for the potential form of _suru_ and _nasaru_ for the polite form of _suru_). Thus (28a) becomes ungrammatical if _suru_ is added to the VN.

(32) *Sakuzitu   syusyoo   ga   [hisaiti : sisatu-si-ta]   sai
       yesterday   prime.minister NOM   [stricken.area : inspect-do-PST]   time
       (Same meaning as 28a)

Needless to say, the exclusion of tense and other inflections is satisfied in post-syntactic compounds derived from noun phrase structure as well. This condition, which may be dubbed "the non-finiteness constraint", constitutes a distinct feature of post-syntactic compounding. The significance of the non-finiteness constraint in Japanese will be increasingly clear as we move on.

Assuming the wordhood of post-syntactic compounds, we now go on to review grammatical restrictions on the compounded nouns. Post-syntactic compounds exhibit much the same range of possible argument-types inside them as do most noun-incorporating verbs in polysynthetic languages discussed by Baker (1988). That is, only the internal argument that is structurally adjacent to the head VN in

clausal or phrasal structure is qualified. This "internal argument" restriction allows three kinds of nouns as targets of post-syntactic compounding: subjects of unaccusative VNs, locative complements of motion VNs, and direct objects of transitive VNs. (33) exemplifies this restriction with typical adverbial clause constructions.

(33) a. Subjects of unaccusative VNs
  tyoonan ga tanzyoo no sai → [*tyoonan* : *tanzyoo*] *no* *sai*
  first.son NOM be.born GEN time [first.son : be.born] GEN time
  'When the first son is/was born'

 b. Goal/locative/source complements of motion VNs
  Syatyoo ga Pari ni tootyaku no sai
  president NOM Paris DAT arrive GEN time

  → *Syatyoo ga *[*Pari* : *tootyaku*] *no* *sai*
   president NOM [Paris : arrive] GEN time
  'When the president arrived/arrives in Paris'

 c. Direct objects of transitive VNs
  Keisatu ga yoogisya o taiho no sai
  police NOM suspect ACC arrest GEN time

  → *Keisatu ga* [*yoogisya* : *taiho*] *no* *sai*
   police NOM [suspect : arrest] GEN time
  'When the police arrested/arrest the suspect'

The application of post-syntactic compounding is completely free of lexical restrictions. Any VN that meets one of the characterizations in (33a), (33b), and (33c) can freely combine with the designated nouns, and native speakers usually do not even notice that compounding is taking place there. The archetypal VNs that pattern with (33a) include *hassei* 'happen', *bakuhatu* 'explode', *tanzyoo* 'be born', *syuuryoo* 'end', and *hunka* 'explode'; those that pattern with (33b) include *nyuugaku* 'enter (school)', *syuttyoo* 'make a business trip', and *taizai* 'stay'; and those that pattern with (33c) include *hoomon* 'visit', *koonyuu* 'purchase', *tyoosa* 'investigate', *kensetu* 'build', *sakusei* 'make', and *hakai* 'destroy'.

 The internal-argument restriction automatically rules out combinations of transitive VNs and their subjects, as shown in (34a), and combinations of any type of VNs and adjuncts, as shown in (34b).

(34) a. *Yoogisya o keisatu ga taiho no sai*
  suspect ACC police NOM arrest GEN time

  → *Yoogisya o [*keisatu* : *taiho*] *no* *sai*
   suspect ACC [police : arrest] GEN time
  (Same meaning as 33c)

b. *Keisatu ga   yoogisya o   ekimae   de   taiho no   sai*
   police   NOM suspect   ACC station.front LOC arrest GEN time

  → **Keisatu ga   yoogisya o   [ekimae : taiho]   no   sai*
     police   NOM suspect   ACC [station.front : arrest] GEN time
  'When the police arrested/arrests the suspect in front of the station'

The exclusion of transitive subjects and adjuncts is parallel to many instances of polysynthetic noun incorporation and appears to constitute strong evidence for the syntactic derivation.

    It should be carefully noted here that the subjects of certain unergative verbs appear to be amenable to post-syntactic compounding, although these subjects are generally looked upon as "external" rather than "internal" arguments. Examples like the one in (35), pointed out to me by Hideki Kishimoto, sound more or less acceptable.

(35) *Hon-daigaku   ni   Yamada-kyoozyu   ga   syukkoo   no   sai*
    this-university DAT Yamada-professor NOM give.lecture GEN time

  → ?*Hon-daigaku   ni   [Yamada-kyoozyu : syukkoo]   no   sai*
     this-university DAT [Yamada-professor : give.lecture] GEN time
  'When Professor Yamada came to this university to give a lecture'

Since the subject 'Professor Yamada' represents a volitional agent, the example in (35) appears to involve incorporation of an external argument in violation of the internal-argument constraint. In view of the fact that archetypical unergative verbs such as *undoo* 'exercise', *katuyaku* 'play an active role', and *zisatu* 'commit suicide' firmly resist compounding with their subjects, however, it is not appropriate to extend the domain of post-syntactic compounding to the subjects of unergative verbs unconditionally. What is intriguing about (35) is that the VN *syukkoo* meaning 'go to a university and give a lecture there (as a part-time lecturer)' involves the notion of motion (i.e. movement to a place), as opposed to *koogi*, which denotes a simple activity of 'giving a lecture' and does not undergo post-syntactic compounding. In fact, *kikoku* 'return to one's home country', *syutuzin* 'go to the war front', and other VNs that involve volitional motion exhibit the same pattern as (35). The relatively high acceptability of (35) will then be attributed to the lexical semantic structures of motion VNs. In terms of Lexical Conceptual Structure, non-agentive motion verbs like *fall* and *roll* are represented simply as [x MOVE PATH], whereas agentive motion verbs like *run* and *walk* are assumed to have a complex structure like [x ACT] CAUSE [x MOVE PATH] (i.e. x moves by x's own volitional action) (Kageyama 1996). The VN *syukkoo* has the latter complex structure, where the syntactic subject (*x*) corresponds to both the agent of action (which is normally linked to the external argument) and the mover (which is normally linked to the internal

argument). The existence of the mover (i.e. internal) argument in lexical semantic structure is considered responsible for the relative acceptability of (35).

The syntactic character of post-syntactic compounds is reinforced by the occurrence of the honorific prefix *go-* on the host VN, as in (36a), and further by the ability of the incorporated noun to participate in sentential anaphora, as in (36b).

(36) a. *Tooten   de   [keitaidenwa : go-koonyuu]   no   okyakusama wa*
        our.store LOC [cellphone : HON-purchase] GEN customers   TOP
        'the customers who purchase cellphones at our store'

    b. *Keisatu wa   [yoogisya$_i$: taiho] no   sai, sono$_i$ tuma mo renkoo-si-ta.*
       police   TOP [suspect$_i$: arrest] GEN time, his$_i$ wife also take-do-PST
       'When the police arrested the suspect, they also took his wife to the police station.'

It has so far been assumed that the VNs that trigger post-syntactic compounding belong to the Sino-Japanese stratum. As observed by Kageyama (1993), however, native VNs are also qualified. By native VNs are meant native (compound) verbs that are converted to nouns and behave as VNs in taking the light verb *suru*. There are two groups of such native VNs. Compare (37) with (38).

(37)   Native VNs directly nominalized from V-V compound verbs
     Tosyokan kara tosyo o   moti-dasi   no   sai
     library   ABL book ACC taking-out GEN time
     'When you take out a book from the library'

    → *Tosyokan   kara   [tosyo : moti-dasi]   no   sai*
        library   ABL   [book : taking-out] GEN time

(38)   Native VNs resulting from compounding of nominalized verbs
     **Syoten   de   manga o   tati-yomi   no   sai*
     bookstore LOC comic ACC standing-reading GEN time
     'When I was browsing a comic book in the bookstore'

    → **Syoten   de   [manga : tati-yomi]   no   sai*
        bookstore   LOC [comic : standing-reading] GEN time

The VN *moti-dasi* 'taking out' in (37) originates from the tensed compound verb *moti-dasu* 'take out'. Since the base verb has the ability to assign accusative marking, the nominalized VN takes over the case-assigning ability and assigns the accusative case to the object 'book' in (37). This case relation makes the post-syntactic compounding possible.

By contrast, the VN *tati-yomi* [standing-reading] 'browsing while standing' does not come directly from a finite compound verb like **tati-yomu*, which is non-existent,

but from the compounded deverbal nouns *tati* 'standing' and *yomi* 'reading'. Since nouns have no case marking ability, the accusative marking in the clause structure of (38) is considered illicit in the first place, and the compound that originates from that structure is naturally ungrammatical.

The validity of the constraints reviewed above gains dramatic confirmation from an entirely new syntactic environment involving native honorific verbs.

(39)   *teiki-ken*   *o*   *o-motome*   *no*   *kata*   *wa*
         commuter-ticket   ACC   HON-purchase   GEN   person   TOP

       →   [*teiki-ken* : *o-motome*]   *no*   *kata*   *wa*
           [commuter-ticket : HON-purchase]   GEN   person   TOP
           'people who want to buy a train pass'

The honorific prefix *o-*, when attached to the *ren'yō* base of native verbs, functions to nominalize the verb while retaining its case-marking property. The construction in (39) thus meets all of the non-finiteness condition, case-marking condition, and internal-argument condition, thereby yielding a well-formed post-syntactic compound.

The involvement of case, a notion which is syntactic par excellence, augments the syntactic nature of post-syntactic compounds. Unlike NI in polysynthetic languages, where incorporated nouns are characteristically indefinite and non-specific (van Geenhoven 1998; Carlson 2006; and many others), the Japanese post-syntactic compounding is insensitive to the definiteness, specificity, or referentiality of compounded nouns, thus allowing even proper nouns. This is part of the reason those compounds are called "post-syntactic".

3.2 N-AN compound predicates

Work by Namiki (1988), Yumoto (1990), and Kageyama (1993) suggests that post-syntactic compounding has been extended from VN heads to AN heads. ANs or adjectival nouns are nouns that have an adjectival function (Chapter 2 [Kishimoto and Uehara, this volume]). Specifically, the characteristic accentuation and the internal-argument restriction hold for post-syntactic compounding with ANs as well.

(40)   a.   *Nihongo*   *ni*   *koyuu*   *no*   *tokutyoo*
          Japanese   DAT   inherent   GEN   feature

         →   [*Nihongo* : *koyuu*]   *no*   *tokutyoo*
            [Japanese : inherent]   GEN   feature
          'features inherent in Japanese'

b.　*singi*　　*ga*　　*huzyuubun*　*ni*　　*tuki*
　　discussion　NOM　inadequate　DAT　due.to

　→　[*singi* : *huzyuubun*]　　　*ni*　　*tuki*
　　　[discussion : inadequate]　DAT　due.to
　'due to the fact that the discussion is inadequate'

In (40a), the AN *koyuu* 'inherent', a two-place predicate taking the argument structure of (Subject (Complement)), incorporates the dative complement that is adjacent to it. In (40b), on the other hand, the AN *huzyuubun* 'inadequate' is a one-place predicate selecting a subject as internal argument, so that the subject noun undergoes incorporation. In contradistinction to these, adjuncts are totally prohibited from incorporation.

(41)　*Kokkai*　*de*　　*singi*　　*ga*　　*huzyuubun*　*ni*　　*tuki*
　　　Diet　　LOC　discussion　NOM　inadequate　DAT　due.to
　　　'due to the fact that the discussion is inadequate in the Diet'

　→　**singi*　　　*ga*　　[*kokkai* : *huzyuubun*]　*ni*　　*tuki*
　　　discussion　NOM　[Diet : inadequate]　　DAT　due.to

Interestingly, *ninki* 'popularity' in (42), which superficially takes the same case pattern as the AN *koyuu* 'inherent' in (40a), cannot trigger post-syntactic compounding (Kageyama 1993).

(42)　*wakamono*　　*ni*　　*ninki*　　*no*　　*myuuzisyan*
　　　young.people　DAT　popularity　GEN　musician
　　　'the musician who is popular among the young'

　→　**[wakamono* : *ninki*]　　　*no*　　*myuuzisyan*
　　　young.people : popularity　GEN　musician

The ungrammaticality of the compound **[wakamono* : *ninki*] is attributed to the lexical status of *ninki* as a simple noun rather than an AN. The dative marker in (42) is not directly assigned by the noun *ninki* 'popularity' but is associated with a concealed verb *aru* 'be' (i.e. *wakamono ni ninki ga aru* [young.people DAT popularity NOM be]). This means that an internal argument can be incorporated only if it is directly case-marked by the host AN – a condition that is shared by post-syntactic compounding with VNs. The requirement of case reinforces the syntactic nature of post-syntactic compounding.

4 Non-finite compound predicates

The observations in Sections 2 and 3, taken together, reveal an intriguing generalization concerning the variation in the NI family. The N-V compound verbs in Section 2.1 and the N-A compound adjectives in Section 2.2, which are subject to a multiplicity

of lexical conditions and have little productivity, are finite with tensed verbs and adjectives in the head, whereas the post-syntactic compounds of the forms "N-VN" and "N-AN" discussed in Section 3, which are fully productive, are non-finite with their heads having a nominal function.

Now, a broader observation suggests that the finite/non-finite distinction is not a dichotomous parameter but forms a continuous cline as depicted in (43), where the native tensed compounds, which are unproductive, occupy one extreme end and the post-syntactic compounds, which are fully productive, occupy the other end.

(43)   Non-finiteness cline for noun incorporation
       less productive   (semi-productive)   more productive
       ←- →
       tensed N-V / N-A   (adnominal and gerundive)   tenseless N-VN / N-AN

The non-dichotomous characterization of the finiteness/non-finiteness parameter is motivated by the discovery of hitherto neglected types of non-finite N-V and N-A compounds belonging to the native stratum that fall between the extremes. They are classified into two groups: adnominal and gerundive.

The native N-V compounds licensed in adnominal environments (relative and appositive clauses) share the composition "[_{compound} N + adnominal V]" followed by head nouns, as exemplified by (44).

(40)   [*miti-yuku*] *hito* [street-go people] 'people walking the streets, passers-by',
       [*kaze-kaoru*] *kisetu* [wind-be.fragrant season] 'the season when fragrant wind
       blows', [*kiki-semaru*] *bamen* [demon-approach situation] 'a ghastly situation',
       [*yuuki-aru*] *koodoo* [courage-have action] 'a brave action', [*kokoro-atatamaru*]
       hanasi [heart-get.warm story] 'a heart-warming story', [*kokoro-yasumaru*] *basyo*
       [heart-feel.rested place] 'a comfortable place', [*baitaritii-ahureru*] *dansei*
       [vitality-be.full man] 'an energetic man', *ten ni* [*tuba-suru*] *kooi* [heaven DAT
       spit-do action] 'an infamous action that will backfire on the person himself',
       kane ga [*mono-yuu*] *sekai* [money NOM thing-tell world] 'the world where only
       money works', [*usirogami-hikareru*] *omoi* [back.hair-pull.PASS feeling] 'a sad
       feeling of parting with someone', [*mizu-nurumu*] *kisetu* [water-get.warmer
       season] 'the season when water gets warmer', [*iki-zumaru*] *sessen* [breath-
       choke close.match] 'a breath-taking close match', [*koke-musita*] *iwa* [moss-
       grew rock] 'a moss-covered rock', [*hone-oreru*] *sigoto* [bone-break task]
       'a painstaking task', [*ti-nurareta*] *katana* [blood-smear.PASS sword] 'a blood-
       smeared sword', *kunsyu ni* [*yumi-hiku*] *kooi* [lord DAT bow-draw action]
       'a rebellious action', [*hana-saku*] *kisetu* [flower-bloom season] 'a season when
       flowers come into bloom', [*kokoro-ubau*] *utagoe* [heart-captivate singing.voice]
       'a captivating singing voice', [*kokoro-hiku*] *egao* [heart-attract smile]
       'an attractive smile', [*kokoro-arawareru*] *omoi* [heart-wash.PASS feeling]
       'a refreshed feeling'

The expressions shown in the brackets are considered to make up a morphological unit of compound because they lack case particles inside and cannot be broken up by interpolation of adverbs and other syntactic elements. Those compound-like expressions are acceptable only if they are used as relative clauses modifying the following head nouns. Specifically, they are unable to conclude a sentence with tense inflections, as demonstrated by the ungrammaticality of (45).

(45) a. *Oozei   no   hito   ga   [miti-it]-ta.   cf. *miti-yuku hito* 'passer-by'
          many   GEN   people   NOM   [street-go]-PST
          'Many people passed by.'

      b. *Sen-getu   wa   [kaze-kaot]-ta.   cf. *kaze-kaoru kisetu*
          last-month   TOP   [wind-be.fragrant]-PST   'season of fragnant wind'
          'Wind was fragrant last month.'

It is instructive to compare the last two examples in (44) – *kokoro-ubau* [heart-captivate] 'captivating' and *kokoro-hiku* [heart-attract] 'attractive' – with their tensed counterparts, which, as mentioned in Section 2, are legitimate only in the passive forms, *kokoro-ubawareru* [heart-captivate.PASS] and *kokoro-hikareru* [heart-attract.PASS]. Note the absence of *rendaku* in *kokoro-**h**iku* [heart-attract], as contrasted with *kokoro-**b**iku* in *Man'yoshu* poetry of the 8th century (Kinuhata 2010: 36). This indicates that the N-V compound verbs appearing in prenominal position with the non-past tense inflection are distinct from those that function to conclude a sentence with tense inflections.

Semantically, the adnominal compound verbs in (44) designate an adjectival concept characterizing the modified noun, as shown by their English translations. The adjectival function is attributed to the verbs' non-past (or adnominal) inflection that represents a constant or permanent state unaffected by the progress of time, or by the past-participial *-ta* that denotes a resultant state of the noun described rather than a particular past time. In other words, the inflections *-ru* and *-ta* in those examples do not represent tense as is the case with ordinary finite clauses and therefore are considered 'non-finite'.

The adnominal restriction holds for compound adjectives as well, as shown in (46).

(46) [*mime-uruwasii*] *zyosei* [appearance-graceful lady] 'a beautiful lady', [*kokoro-nai*] *kotoba* [consideration-null word] 'inconsiderate', [*kitan-nai*] *iken* [restraint-null opinion] 'a candid opinion', [*toritome-nai*] *hanasi* [point-null story] 'a pointless story', [*yondokoro-nai*] *zizyoo* [support-null situation] 'an unavoidable situation', [*yangoto-nai*] *okata* [disregard-null person] 'a noble person', [*yondokoro-nai*] *riyuu* [anchor-null reason] 'an unavoidable reason'

Particularly productive is a class of compound adjectives headed by the negative *nai* 'be null, be absent, be lacking', such as *hirui-nai N* [comparable.thing-null] 'incomparable, unparalleled' and *yoosya-nai N* [mercy-null] 'merciless'. These compound adjectives will gain increased formality if *nai* is substituted for by the classical adnominal inflection *naki*, and in fact there are examples that are fixed with the classical *naki*, such as *ie-naki (ko)* [home-null] 'homeless (children)', *aku-naki (doryoku)* [satisfaction-null] 'insatiate, unceasing (efforts)', *owari-naki (tabi)* [end-null] 'endless (journey)', *zingi-naki (tatakai)* [moral.code-null] 'inhumane (fight)', *kokkyoo-naki (isidan)* [border-null] 'borderless (doctors)', and likewise *X-naki (ato)* [X-die] '(after) X passed/passes away'.

The second group of non-finite N-V compounds shares the gerundive form in the V elements. One of the syntactic functions of the gerundive is to designate adverbial meanings corresponding to converbs in many languages, as in *arui-te* [walk-GER] 'on foot'. The gerundive compound verbs in (47) also bear such an adverbial function.

(47) *tegusune-hii-te* [hand.ointment-apply-GER] 'being ready (to shoot an arrow), waiting eagerly', *asemizu-tarasi-te* [sweat-drip-GER] 'very hard', *kokoro-si-te* [attention-pay-GER] 'with great care', *hitai ni ase-site* [forehead DAT sweat-do-GER] '(work) very hard', *oto-tate-te* [noise-make-GER] 'with big sounds', *na-zasi-te* [name-point-GER] 'by explicit mention of someone's name'

A negative form (*-zu*) is also used in this type of compound, as in *namikaze-tate-zu ni* [wave.wind-make-NEG DAT] 'without arousing trouble, calmly'. These examples lose acceptability if used as the tensed main verb of a sentence.

Gerundive non-finite compounds are attested with adjectives as well. The N-A compounds in (48), where *-ku* is the marker for the gerundive of adjectives, all have an adverbial function.

(48) *asi-sige-ku* [leg-frequent-GER] '(visit) frequently', *tei-yo-ku* [appearance-good-GER] 'tactfully', *kokoro-yo-ku* [heart-good-GER] 'willingly', *saigen-na-ku* [end-null-GER] 'endlessly', *monku-na-ku* [dispute-null-GER] 'undisputedly', *enryo-na-ku* [reserve-null-GER] 'unreservedly', *haba-hiro-ku* [width-broad-GER] 'widely', *baransu-yo-ku* [balance-good-GER] 'in a balanced manner', *kigen-yo-ku* [temper-good-GER] 'in a good temper', *kimoti-yo-ku* [feeling-good-GER] 'pleasantly', *aisoo-yo-ku* [friendliness-good-GER] 'amiably', *ikioi-yo-ku* [force-good-GER] 'vigorously', *kooritu-yo-ku* [efficiency-good-GER] 'efficiently', *mibae-yo-ku* [appearance-good-GER] 'attractively', *te-atu-ku* [hospitality-rich-GER] 'hospitably', *yogi-na-ku* [alternative-null-GER] 'unavoidably'

Particularly high productivity is observed with such head adjectives as *-yo-ku* 'good' and *-na-ku* 'without'. The acceptability of these gerundive compound adjectives is also degraded if they are put in the tensed predicate of a sentence.

In addition to the compound verbs and compound adjectives exemplified above, miscellaneous N-X compounds fall under the rubric of non-finite compounds.

(49)   a.   N-mimetic
           yaruki-manman [eagerness-overflowing] 'very eagerly', *nekki-munmun* [heat-steaming] 'filled with excitement'

       b.   N-N *ni*
           syorui-katate ni [paper-one.hand DAT] 'with papers in the hand', *kankookyaku-aite ni* [tourist-target DAT] 'targeting at tourists', *kane-meate ni* [money-aim DAT] 'for money'

       c.   N-deverbal N (Kageyama 1993)
           [*Karayan-nakiato*] *no ongakukai* [[Karajan-die.after] GEN music.scene] 'the music scene after the death of Karajan', [*Yul Brynner-katanasi*] *no sukin-heddo* [[Yul Brynner-be.ruined] GEN skin-head] 'the skinhead that will make Yul Brynner lose face'

Like nouns, mimetics do not inflect for tense and hence are considered non-finite.

To summarize, this section has revealed that there is a huge variety of non-finite compound predicates including adnominal compound verbs, gerundive compound verbs and adjectives, and others. In the previous literature, these miscellaneous kinds of N-Pred compounds that are legitimate only in non-finite environments have gone totally unnoticed. The non-finiteness cline proposed in (43) not only gives due recognition to them, but also accounts for their semi-productive creation properly by situating them between the non-productive tensed N-V/N-A compounds in Section 2 and the fully productive post-syntactic N-VN/N-AN compounds in Section 3 on the non-finiteness cline (43).

5 Agent compounding

All the N-incorporation and N-compounding processes discussed so far neatly fall under the supposedly universal constraint on argument incorporation that stipulates that incorporation and compounding take place on the internal argument of a predicate. This constraint is further ascribed to the universal principle of argument realization (Levin and Rappaport Hovav 2005) that determines the correspondences between semantic roles and grammatical relations, as schematically shown in (50).

(50)   a.   Canonical argument realization of the transitive verb "kill"
           semantics: Agent ('killer')   Theme ('killee')
                              ↓   ↓
           syntax: External argument (Subject)   Internal argument (Object)

In point of fact, previous literature on noun incorporation in polysynthetic languages has attested no real example in which an agent is incorporated with a transitive verb. It has been found, however, that Japanese has a productive way of incorporating or compounding an agent argument with a transitive VN. Such a situation is certainly inconceivable with ordinary noun incorporation, as shown by the contrast between (33c) and (34a), reproduced below as (51a) and (51b), respectively.

(51) a. *Keisatu   ga   yoogisya   o   taiho   no   sai*
          police   NOM   suspect   ACC   arrest   GEN   time

          → *Keisatu   ga   [yoogisya : taiho]   no   sai*
              police   NOM   [suspect : arrest]   GEN   time
          'When the police arrested/arrests the suspect'

     b. *Yoogisya   o   keisatu   ga   taiho   no   sai*
          suspect   ACC   police   NOM   arrest   GEN   time

          → **Yoogisya   o   [keisatu : taiho]   no   sai*
              Suspect   ACC   [police : arrest]   GEN   time

The ban on agent incorporation is thus deemed a truly universal principle that applies to predicational structure. When we shift attention to modificational structure, however, we find that this canonical rule is violated. Observe the examples in (52). The symbol ":" inside the brackets indicates the same kind of phonological break that is observed with post-syntactic compounds. (In (52) and subsequent examples, *no* is glossed as 'genitive', but it might be better identified as the prenominal variant of the copula *da* 'be'.)

(52) a. [*Spielberg-kantoku : seisaku*]   *no*   *eiga*
          [Spielberg-director : make]   GEN   movie
          'a movie that Spielberg directed'
          N.B. 'Spielberg' = agent, 'movie' = theme

     b. [*zyosei-pairotto : soozyuu*]   *no*   *hikooki*
          [female-pilot : operate]   GEN   plane
          'the plane that a female pilot flew'
          N.B. 'a female pilot' = agent, 'a plane' = theme

One might be tempted to suggest that some kind of invisible passive is responsible for the creating of those examples, just as lexical passive is involved in English compounds like *Spielberg-produced (films)*. There is every reason to believe that this is not the case.

First, English passive compounds may take not only an agent but also an adjunct as the first member, as in *machine-made* (instrumental) and *Paris-based*

(locative), whereas Japanese agent compounds reject adjuncts, thus excluding compounds like *[kikai : seisaku] no N* 'machine-made N' and *[amerika : kyooiku] no N* 'American-trained N'.

Second, as is the case with post-syntactic compounds, case markers may be realized on the nouns.

(53) a. [*zyosei-pairotto* {***ga / *niyoru***} *soozyuu*] *no* *hikooki*
      female-pilot   {NOM / *by}   operate] GEN plane
      'the plane a female pilot operated'

   b. [*Spielberg* {***ga / *niyoru***} *seisaku*] *no* *eiga*
      [Spielberg {NOM/*by}   make]   GEN movie
      'a movie Spielberg directed'

The agent nouns can be optionally marked with the nominative *ga*, but not with the agent marker *niyoru* 'by'.

Another indicator of the subjecthood of the incorporated agent nouns is the optional appearance of subject honorification on the VN. The subject honorification marker may show up only when the target of honorification is realized as the grammatical subject, and not as a passive agent.

(54) a. [*Yamada-kyoozyu* : ***go****-sippitu*]   *no*   *ronbun*
      Yamada-professor : HON-write.article] GEN article
      'the article Prof. Yamda wrote'

   b. [*o-kyaku-sama* : ***go****-sitei*]   *no*   *ginkoo-kooza*
      HON-customer : HON-designate] GEN bank-account
      'the bank account which our customers designate'

If passive is irrelevant to agent compounds, how can such compounds that deviate from the canonical rule of argument realization be accounted for? The answer proposed by Kageyama (2009, 2013) hinges on the distinction of two types of predication: event (or stage-level) predication vs. property (or individual-level) predication. The former represents an event or temporary state that unfolds according as time progresses, and the latter a more-or-less permanent attribute characteristic of the subject entity. In the literature on event semantics (cf. Krifka et al. 1995), it is commonly known that event predication is compatible with punctual time adverbials but property predication is not, as illustrated in (55a) and (55b).

(55) a. Firemen were not available at that time. (Stage-level)

   b. Firemen are altruistic (*at this moment). (Individual-level)

The agent compounds under discussion behave as property predications, excluding temporal adverbials referring to a particular time, as in (56), and expressions indicating the progress of an event, as in (57).

(56) ***kinoo   no   yoru   ni*** [*zyosei-pairotto* :   *soozyuu*]   *no*   *hikooki*
  yesterday GEN   night   DAT [female-pilot :   operate]   GEN   plane
  'the plane that a female pilot flew last night'

(57) *[*zyosei-pairotto* :   *soozyuu*]-***tyuu***   *no*   *hikooki*
  [female-pilot :   operate]-in.the.middle   GEN   plane
  'the plane that a female pilot is flying'

Note that post-syntactic compounds, which are event predications, freely allow temporal adverbials and the suffix -*tyuu* 'in the middle of'.

 The aberrant argument realization in agent-incorporating compounds should then be attributed to their semantic function of property predication, as opposed to event predication in post-syntactic and all other kinds of compounds. The property predication function makes it possible to characterize a noteworthy property of the noun entity (i.e. the nouns modified by agent compounds). Because of the noteworthiness requirement, the incorporated agents must be embodied by nouns that are informative enough to characterize the notable property of the modified nouns. Consider the alternative nouns in the agent position in the following examples.

(58) a. [{*zyosei-pairotto/*pairotto/*Ø*}:   *soozyuu*]   *no*   *hikooki*
   [{female-pilot/*pilot/*Ø} :   operate]   GEN   plane
   'the plane flown by {a female pilot/a pilot/Ø}'

  b. [{*Spielberg/*hito/*Ø*} :   *seisaku*]   *no*   *eiga*
   [{Spielberg/*person/*Ø} :   produce]   GEN   film
   'the film produced by {Spielberg/a person/Ø}'

The nouns 'pilot' and 'human' are rejected in (58a) and (58b), respectively, because they are not informative enough. Both examples naturally preclude a zero noun.

 While the preceding examples involve Sino-Japanese VNs, an analogous case can be made with native verbs in adnominal positions. Compare the examples in (59) involving agents inside the compounds with the adnominal N-V compounds discussed in the previous section that involve internal arguments.

(59) a. [*Karayan* :   *hikiiru*]   *Berurin*   *firuhaamonii*
   [Karajan :   lead]   Berlin   Philharmonic
   'the Berlin Philharmonic Orchestra led by Karajan'

b. [*Robaato de Niiro* : *enziru*]   *Aru   Kapone*
   [Robert De Niro :   perform]   Al   Capone
   'the role of Al Capone performed by Robert De Niro'

The native verbs qualified for this construction are very limited.

From the foregoing discussions, it can be safely concluded that agent compounds are exempt from the canonical argument realization pattern and the universal ban on agent incorporation, which are pertinent only to stage-level or event predications. Agent compounds belong to a distinct predication type, i.e. property or individual-level predication, in which the canonical rules of argument realization often fail to apply. To give only one example, the transitive verb *kill* must realize its object obligatorily in event predication sentences, as in (60a), but it can omit the object in property predication sentences, as in (60b).

(60) a. Event predication: The tiger killed *(the bear) last night.

b. Property predication: Tigers kill only at night. (Goldberg 2001)

The deviation from the canonical argument realization appears to be a prerogative of property predication sentences across the world's languages (see Kageyama 2006, 2009 for concrete examples).

6 Conclusion and future research perspectives

This chapter has surveyed the basic properties of N-V, N-A, N-VN, and N-AN compounds available in contemporary Japanese and discovered parametric variation on the finiteness/non-finiteness cline. The heads of N-V and N-A compounds belong to the native stratum and have only a low degree of productivity, whereas those of N-VN and N-AN compounds basically come from the Sino-Japanese vocabulary and have high productivity. A novel finding of this paper is the existence of non-finite N-V and N-A compounds which exhibit varying degrees of productivity in limited syntactic environments of adnominal positions and gerundive inflections. Agent compounds, which depict the individual-level property of their modified entity, occupy the extreme end of non-finiteness. Undoubtedly, the high productivity of N-VN, N-AN, and agent compounds is further tied in with the virtually unlimited productivity of compound nouns including deverbal (synthetic) compound nouns such as *kane-mooke* 'moneymaking'.

These observations raise new issues for future studies of Japanese. First, why are N-compounding and N-incorporation constrained by non-finiteness? Second, why do N-compounding and N-incorporation as a whole differ markedly in productivity from V-compounding and V-incorporation, discussed in Chapter 8 (Kageyama, this volume),

which are productive in finite clauses? A possible answer would be to say that V-compounding and V-incorporation are productive because the agglutinative character of the language that allows one verb to be annexed to another iteratively. This answer, however, begs the question of why the agglutinative character applies only to the verbal domain and not to the nominal domain.

The Japanese data also raises a new problem for the typology of noun incorporation. Based on a rich array of data on Noun Incorporation in various languages (Mithun 1984; Sadock 1985; Baker 1988), Rosen (1989) proposed to divide NI into three types, depending on the argument structural properties.

A. Compounding type: Polynesian (Samoan, Tongan, Niuean, etc.), Micronesian (Mokilese, Ponapean, Kusaiean, etc.). In this type, the object is incorporated with the transitive verb, so that the whole verb complex behaves as an intransitive, as in (61). The incorporated noun is indefinite and non-referential, and no stranding or doubling is possible.

(61) *Ngoah* [*ko   oaring*]$_V$.
      I   grind-coconut
      'I am coconut-grinding.'

B. Classifier type: Northern Iroquoian languages (Mohawk, Seneca, Oneida, nondaga, Cayuga, Tuscarora), Caddo, Rembarnga. This type allows object doubling, as schematically represented by English words in (62).

(62) a. 'I animal-caught a dog.'

      b. 'I chicken-eat that chicken.'

      c. 'I chicken-eat that (*pro*).'

Here an object that is semantically related to the incorporated noun can be realized syntactically. Specificity and referentiality are determined by the syntactic object showing up outside the verb complex. Since doubling of the incorporated object is allowed as in (62b), the example in (62c), which at first blush appears to be a case of stranding of the demonstrative, is analyzed as having an invisible pronoun (*pro*) as the object. So (62c) is looked upon as a case of doubling on a par with (62b).

C. Stranding type: Southern Tiwa, Greenlandic Eskimo. Here, stranding of a determiner is possible, as in (63a), but object doubling is said to be disallowed, as in (63b). Because of this discrepancy, it is implausible to hypothesize an invisible pronoun for (63a).

(63) a. 'I chicken-eat that.'

      b. *'I chicken-eat that chicken.'

Since it is impossible to generate only a determiner by a syntactic rule, it is only for this type of language, Rosen contends, that syntactic movement (i.e. Incorporation) of the object noun from the syntactic object position into the verb complex is properly motivated. And in later work, Baker (1996, 2009) presented some other languages of this type.

Rosen (1989) assumed the three types of NI as parameters of language typology. This view, however, is at odds with the two-morpheme Sino-Japanese VNs. As regards how an object argument is realized externally, these compounds are divided into the three types that Rosen set up for distinguishing language types.

(64) Compounding type
    (*hon o) [doku-syo]-suru   [(*book ACC)   [read-book]-do]
    (*yuusyoku o) [syoku-zi]-suru [(*dinner ACC) [eat-thing]-do]

(65) Classifier type
    kinzyo no byooin ni [nyuu-in]-suru [nearby GEN hospital DAT [enter-hospital]-do]
    gakuhi o [soo-kin]-suru   [tuition.fee ACC [send-money]-do]

(66) Stranding type
    ude o [kos-setu]-suru   [arm ACC [bone-break]-do]
    eberesuto ni [too-tyoo]-suru   [Mt. Everest DAT [climb-top]-do]

Since the variety of argument realizations that Rosen hypothesized as parameters of language typology are observed within a single language, they should better be regarded as variation on lexical rules rather than as language types.

Acknowledgments

References

Baker, Mark. 1988. *Incorporation: A theory of grammatical relation changing*. Chicago: University of Chicago Press.
Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.
Baker, Mark. 2009. Is head movement still needed for noun incorporation? *Lingua* 119. 148–165.
Carlson, Greg. 2006. The meaningful bounds of incorporation. In Svetlana Vogeleer and Liliane Tasmowski (eds.), *Non definiteness and plurality*, 35–50. Amsterdam: John Benjamins.

Fabb, Nigel. 1998. Compounding. In Andrew Spencer and Arnold Zwicky (eds.), *The handbook of morphology*, 66–83. Oxford: Blackwell.

Gerdts, Donna B. 1998. Incorporation. In Andrew Spencer and Arnold Zwicky (eds.), *The handbook of morphology*, 84–104. Oxford: Blackwell.

Goldberg, Adele E. 2001. Patient arguments of causative verbs can be omitted. *Language Sciences* 23. 503–524.

Hale, Ken and Samuel J. Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Kageyama, Taro. 1980. *Goi no kōzō* [Lexical structures]. Tokyo: Shōhakusha.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2006. Property description as a voice phenomenon. In Tasaku Tsunoda and Taro Kageyama (eds.) *Voice and grammatical relations*, 85–114. Amsterdam: John Benjamins.

Kageyama, Taro. 2009. Structural constraints and predication functions in language. *Gengo Kenkyu* 136. 1–34.

Kageyama, Taro. 2013. Post-syntactic compounds and semantic head-marking in Japanese. In Bjarke Frellesvig and Peter Sells (eds.), *Japanese/Korean linguistics 20*, 363–382. Stanford: CSLI Publications.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Lexical strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Masayoshi Shibatani. 1989. Meishiku kara no fukugōgo keisei [Compound formation from noun phrases]. In Susumu Kuno and Masayoshi Shibatani (eds.) *Nihongogaku no shintenkai* [New developments of Japanese linguistics], 139–166. Tokyo: Kurosio Publishers.

Kinuhata, Tomohide. 2010. Jōdaigo no meishi-hōgō nituite [On noun incorporation in Old Japanese]. *Gobun* 92/93. 34–44. Osaka University.

Kishimoto, Hideki and Geert Booij. 2014. Complex negative adjectives in Japanese: The relation between syntactic and morphological constructions. *Word Structure* 7. 55–87.

Kishimoto, Hideki and Satoshi Uehara. this volume. Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Krifka, Manfred, Francis Pelletier, Gregory Carlson, Alice ter Meulen, Godehard Link and Gennaro Chierchia. 1995. *Genericity: An introduction*. In Greg Carlson and Francis J. Pelletier (eds.), *The generic book*, 1–124. Chicago: University of Chicago Press.

Levin, Beth and Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.

Lieber, Rochelle. 1983. Argument linking and compounds in English. *Linguistic Inquiry* 14. 251–286.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60. 847–894.

Namiki, Takayasu. 1988. Fukugōgo no nichi-ei taishō [Comparison of English and Japanese compound words]. *Nihongogaku* 7(5). 68–78. Tokyo: Meiji Shoin.

Nihon kokugo daijiten [The grand dictionary of the Japanese language] 2000–2002. 2nd edition. Tokyo: Shōgakukan.

Roeper, Thomas and Muffy E.A. Siegel. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9. 199–260.

Rosen, Sara. 1989. Two types of noun incorporation: A lexical analysis. *Language* 65. 294–317.

Sadock, Jerrold. 1985. Autolexical syntax: A proposal for the treatment of noun incorporation and similar phenomena. *Natural Language & Linguistic Theory* 3. 379–440.

Sakakura, Atsuyoshi. 1966. *Gokōsei no kenkyū* [A study of word composition]. Tokyo: Kadokawa Shoten.

Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* 13. 250–282.

Selkirk, Elizabeth O. 1982. The syntax of words. Cambridge, MA: MIT Press.

Shibatani, Masayoshi. 1978. Mikami Akira and the notion of 'subject' in Japanese grammar. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics*, 52–67. Tokyo: Kaitakusha.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar: Postsyntactic compounds in Japanese. *Language* 64. 451–484.

Spencer, Andrew. 1991. *Morphological theory: An introduction to word structure in generative grammar*. London: Basil Blackwell.

van Geenhoven, Veerle. 1998. *Semantic incorporation and indefinite descriptions: Semantic and syntactic aspects of noun incorporation in West Greenlandic*. Stanford: CSLI Publications.

Watanabe, Akira. 1999. Remarks on head movement within VP shell. *Proceedings of the Nanzan GLOW 1999*, 461–485. Nagoya: Nanzan University.

Yumoto, Yoko. 1990. Nichi-ei taishō fukugō-keiyōshi no kōzō [The structure of compound adjectives: Comparison of Japanese and English]. *Gengobunka Kenkyū* [Language and Culture] 16. 353–370. Toyonaka: Osaka University.

Yumoto, Yoko. this volume. Chapter 9: Conversion and deverbal compound nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Taro Kageyama
8 Verb-compounding and verb-incorporation

1 Introduction

Japanese has a variety of compound and complex predicate formations, of which Verb-Verb (V-V) compounding is perhaps the most intriguing from theoretical, descriptive, and typological perspectives. Conspicuously absent from European languages, compound and complex verbs consisting of two adjoined verbs are said to be an areal-typological feature of languages of Asia (Masica 1976), and among the Asian languages, Japanese appears to have the richest system of V-V compound verbs, straddling the boundaries of morphology, semantics, and syntax. This chapter will survey the essential properties of V-V compound verbs, with cursory reference to Verb-Adjective compound adjectives when appropriate.

V-V compound verbs are characterized as a morphological amalgamation of two native verbs, where a first verb (V1), in what is traditionally called the *ren'yō* form (lit. 'adverbal' or 'predicate-modifying' form), is morphologically adjoined to the stem of a second verb (V2) and the whole V1-V2 combination inflects for tense. Following Bloch (1946) and Martin (1975), the *ren'yō* form will be referred to as "infinitive" in this chapter. Formally, the infinitive is identical to the verb stem if it ends in a vowel, as in *tasuke* 'help' < stem *tasuke-* 'to help' or *i* 'stay' (as in *i-basyo* 'whereabouts') < stem *i-* 'to be, stay'; otherwise, it takes the shape of a verb stem supplemented with the vowel *i* at the end, as in *kaki* 'write' < stem *kak-* 'to write' and *ari* 'be' (as in *ari-ka* 'place where something is kept') < stem *ar-* 'to be'. Originally, its primary function was to form a subordinate or coordinate predicate used with a following main verb in sentential structure (Frellesvig 2010: 56). While the subordinating (converb) function in syntactic constructions has been taken over by the *-te* gerundive in contemporary Japanese, it is preserved inside morphological structures where the gerundive marker is prohibited due to its syntactic character. Typical examples are shown in (1a, b, c, d), where the infinitives in boldface modify the following elements.

(1) a. V-V compound verbs: ***tataki****-tubusu* [strike-smash] 'knock to pieces', ***tabe****-kuraberu* [eat-compare] 'compare the taste (by eating)'

      b. V-N compound nouns: ***yaki****-niku* [broil-meat] 'broiled meat', ***nomi****-mizu* [drink-water] 'drinking water'

      c. V-A compound adjectives: ***musi****-atui* [steam-hot] 'steaming hot', ***koge****-kusai* [burn-smelly] 'smell of something burning'

d. Deverbal nouns with suffixes: **_utai_**-_te_ [sing-suffix] 'singer', **_yomi_**-_kata_ [read-suffix] 'how to read'

e. V-to-N conversion (nominalization): **_tasuke_** 'help', **_asobi_** 'play'

f. Compounds of converted nouns: _naname_-**_yomi_** [diagonal-reading] 'cursory reading', _hito_-**_dasuke_** [person-help] 'a kind act, a help'

In addition, infinitives have a secondary usage of V-to-N conversion or nominalization as shown in (1e) and (1f), which presumably stems from the original subordinating function (Frellesvig 2010: 57).

The dual function of subordination and nominalization gives rise to two superficially similar types of compounds, V-V compound verbs of type (1a) and deverbal compound nouns of type (1f), which can be unequivocally differentiated by the way they inflect. V-V compound verbs directly take tense inflection as in _[tabe-kurabe]-ta_ ([eat-compare]-PST), whereas deverbal compound nouns require the support of the light verb _suru_ 'do' to represent tense, as in _[naname-yomi]-si-ta_ ([diagonal-reading]-do-PST). There is also a strong tendency that V-V compound verbs seldom show _rendaku_ or sequential voicing on the initial voiceless obstruent of the second member of a compound, which regularly takes place in compound nouns (Vance 1987), as shown by the contrast between the compound verb _asobi-**t**ukareru_ [play-get.tired] 'get tired from playing' (voiceless [ts] on the second verb) and the deverbal compound noun _asobi-**z**ukare_ [play-fatigue] 'fatigue after playing' (voiced [dz] on the second noun). Some V-V compound verbs that exhibit _rendaku_, as in _ike-**d**oru_ 'catch alive' and _notare-**z**inu_ 'die by the roadside', are arguably attributed to backformation from compound nouns that already have _rendaku_ in them (Matsumoto 1998).

The complex properties of V-V compound verbs have sparked numerous studies in connection with the morphology-syntax-semantics interface. Surveying major works in the literature, this chapter describes the overall system of V-V compounding in a step-by-step manner. After confirming the word status of V-V compounds, Section 2 introduces the fundamental distinction of lexical versus syntactic V-V compound verbs (Kageyama 1989, 1993), which serves as the cornerstone for subsequent developments in theoretical discussions. Section 3 will probe into the central issues surrounding syntactic compound verbs, and Section 4 into those involving lexical compound verbs. Section 5 summarizes the data surveyed and suggests how they should be approached in future research.

2 Lexical vs. syntactic V-V compounds

This section will first show that V-V compound verbs qualify as "words" (Section 2.1), and then introduce the basic distinction of lexical vs. syntactic compounds (Section 2.2).

2.1 Morphological wordhood

In an agglutinative language like Japanese, where verbal elements appear in a sequence one after another, it is not immediately clear whether a given string of verbs constitutes a word or not. While Nishiyama (1998) treats certain types of V-V compounds as being derived from syntactic serial verb constructions, and in fact, the V-V sequences in Old Japanese appear to be more like serial verbs than morphological compounds (Frellesvig et al. 2010), the wordhood of compound verbs in contemporary Japanese is illuminated by comparing them with the 'V-*te* V' complex verbs like *yon-de simau* [read-GER put.away] 'finish reading', where the gerundive ending *-te* serves as a hallmark of syntactic concatenation (Chapter 11 [Nakatani, this volume]).

Several diagnostics for wordhood are constructed on the basis of the general principle of lexical integrity which stipulates that the internal structure of a word cannot be disrupted by syntactic elements (Chapter 14 [Kageyama, this volume]). A popular test is particle insertion, which examines whether a given sequence of morphemes can be separated by focus particles like *sae* 'even', *mo* 'also', and *wa* (topic). This interruptibility test correctly diagnoses combinations of an infinitive verb with an inflected verb like those in (2) as words, while at the same time excluding 'V-*te* V' concatenations like those in (3) from the domain of compounds.

(2) a. *oyogi-tuku* [swim-arrive] 'swim to' (lexical V-V compound) vs. *oyogi wa tuku [swim TOP arrive]

   b. *tabe-hazimeru* [eat-begin] 'begin to eat' (syntactic V-V compound) vs. *tabe mo hazimeru [eat also begin]

(3) a. *yon-de simau* [read-GER put.away] 'finish reading' vs. *yon-de sae simau* [read-GER even put.away] 'finish even reading'

   b. *tabe-te miru* [eat-GER see] 'try eating' vs. *tabe-te wa miru* [eat-GER TOP see]

There are a handful of exceptional gerundive complexes that are frozen as tight units and reject interpolation of particles, as exemplified in (4).

(4) *mi-te toru* [see-GER take] 'grasp, understand', *kut-te kakaru* [eat-GER stick.to] 'lash out at', *hut-te waku* [fall-GER boil] '(bad thing) occur unexpectedly', *yot-te tatu* [be.based-GER stand] 'rely on', *kit-te suteru* [cut-GER throw.away] 'dismiss'

Because of their specialized meanings, these expressions are regarded as being lexicalized and memorized as such.

The morphological word status of compound verbs is confirmed by subject honorification. Applying at syntactic structure when the speaker wishes to show

deference to the person in subject, the honorification rule sandwiches an infinitive verb with a set of discontinuous morphemes: the honorific prefix *o-* on the left and the verbal coda *ni naru* (lit. 'DAT become') on the right, as in *o-warai **ni naru*** [HON-laugh DAT become] '(A respected person) laughs' (Harada 1976). Since syntactic rules cannot refer to the internal structure of a word, the honorific morpheme should apply to a whole compound verb, as in *o*-[V1-V2] *ni naru*. This is indeed the case with the grammatical examples in (5), which present a sharp contrast to the totally ungrammatical V-*te* V complex in (6).[1]

(5) a. *Sensei   wa   wain   o   o*-[*nomi-kurabe*]   ***ni   nat-ta***. (lexical compound)
        teacher TOP wine ACC HON-[drink-compare] DAT become-PST
        'The teacher compared the taste of different kinds of wine.'

   b. *Sensei   wa   wain   o   o*-[*nomi-hazime*]   ***ni   nat-ta***. (syntactic compound)
        teacher TOP wine ACC HON-[drink-begin] DAT become-PST
        'The teacher began to drink wine.'

(6) **Sensei   wa   wain   o   zenbu o*-[*non-de   simai*]   ***ni   nat-ta***.
     teacher TOP wine ACC   all   HON-[drink-GER put.away] DAT become-PST
     'The teacher finished drinking all the wine.'

The word status of V-V compound verbs and the phrasal status of V-*te* V complexes are corroborated by applying the honorific morpheme only to the second verb. Observe that the grammaticality judgments in (5) and (6) are reversed in (7) and (8).

(7) a. **Sensei   wa   wain   o   nomi   **o-kurabe***
          teacher   TOP   wine   ACC   drink   HON-compare

          ni   nat-ta.   (lexical compound)
          DAT   become-PST
          'The teacher compared the taste of different kinds of wines.'

    b. **Sensei   wa   wain   o   nomi   **o-sugi***
          teacher   TOP   wine   ACC   drink   HON-exceed

          ni   nat-ta.   (syntactic compound)
          DAT   become-PST
          'The teacher drank too much wine.'

---

1 For syntactic compound verbs, only V1 may be sandwiched by the discontinuous honorific morpheme, as in ***o nomi ni nari*** *hazimeru* [HON drink DAT become begin], where *o nomi ni nari* forms part of the complement clause of *hazimeru* 'begin' and the verb *nari* 'become' is morphologically adjoined to *hazimeru*, yielding a structure like [*o nomi ni* t_i] [$_V$ *nari$_i$ hazimeru*].

(8) *Sensei   wa   wain   o   zenbu   non-de   **o-simai   ni   nat-ta**.*
teacher TOP wine ACC all   drink-GER HON-put.away DAT become-PST
'The teacher drank up all the wine.'

Due to its syntactic nature, the honorific prefix *o-* is generally incapable of occurring inside word structures. The examples of (7a) and (7b) are ungrammatical because the morphological sequences *nomi-kuraberu* [drink-compare] and *nomi-sugiru* [drink-exceed] are interrupted by the syntactic honorific prefix that is attached to their second members. The V-*te* V complex of (8), on the other hand, can accommodate the honorific prefix on V2 because the gerunditive *-te* in V1 marks the closure of a phrase.

There are several other diagnoses to ascertain the word status of V-V compound verbs, among which are backward gapping (Kageyama 1989, 1993) and *-kata* 'way' nominalization (Kageyama 1993; Kishimoto 2006). It should be stressed that all these morphological diagnoses produce a solid and uniform result in identifying V-V sequences as morphological words regardless of the distinction between "lexical" and "syntactic" compounds, to be explained in the next subsection.

2.2 Syntactic and lexical V-V compound verbs

This subsection introduces the fundamental distinction of lexical and syntactic compounds in V-V compounding, espoused by Kageyama (1989, 1993). The two classes of compound verbs both make up a word and cannot be distinguished by morphological shape or by accentuation. Observe the examples in (9) and (10).

(9) Lexical V-V compound verbs
uti-korosu [shoot-kill] 'shoot to death', *nagare-otiru* [flow-fall] 'flow down', *aruki-tukareru* [walk-get.tired] 'get tired from walking', *tobi-agaru* [jump-go.up] 'jump up', *naki-sakebu* [cry-shout] 'cry and scream', *huri-yamu* [fall-stop] '(rain) stop falling', *sawagi-tateru* [make.a.fuss-do.excessively] 'make a big fuss'

(10) Syntactic V-V compound verbs
kaki-hazimeru [write-begin] 'begin to write', *tabe-oeru* [eat-finish] 'finish eating', *huri-kakeru* [rain-be.about.to] 'be about to rain', *home-au* [praise-join] 'praise each other', *mi-sobireru* [see-miss] 'miss seeing', *tabe-sugiru* [eat-exceed] 'eat excessively', *ii-wasureru* [say-forget] 'forget to say', *nomi-tukusu* [drink-exhaust] 'drink up'

In traditional Japanese grammar, although lexical compound verbs are indeed treated as compounds, syntactic compound verbs are usually regarded as syntactic verb-auxiliary concatenations. In referring to verbs like *hazimeru* 'begin', *oeru* 'finish',

and *tuzukeru* 'continue' as "aspectual verbs", early generative studies such as Shibatani (1973) and Kuno (1983) also looked upon them as syntactic constructions, and their status as compound verbs, as demonstrated in Section 2.1, was not clearly discerned until Kageyama (1989, 1993). The vital difference between the two classes resides in the syntactic behavior of their first members. Given the non-interruptibility principle of lexical integrity (Chapter 14 [Kageyama, this volume]), if a compound allows a syntactically motivated element to occur inside it (especially in its V1 position), it is deemed syntactic; otherwise, lexical. Capitalizing on this principle, Kageyama (1989, 1993) formulated several criteria for differentiating syntactic from lexical compounds, such as the occurrence or nonoccurrence of passives, causatives, honorific verbs, verbal anaphora, light verb constructions, and idioms in the V1 position of a compound verb, as summarized in Table 1.

Table 1: Syntactic criteria for the lexical-syntactic distinction

	Lexical compound verbs	Syntactic compound verbs
I. Passives in V1	*[*os are*] *aku* [push-PASS]-open(vi.) 'open (vi.) by being pushed'	[*ais are*] *tuzukeru* [love-PASS]-continue 'continue to be loved'
II. Honorific verbs in V1	*[*o uke ni nari*] *toru* [HON-get DAT beecome]-take 'receive'	[*o utai ni nari*] *hazimeru* [HON-sing DAT become]-begin 'begin to sing'
III. Verbal anaphora in V1	*naki sakebu* → *[*soo si*] *sakebu* cry-scream [so do]-scream lit. 'scream by doing so'	*kaki oeru* → [*soo si*] *oeru* write-finish [so do]-finish 'finish doing so'
IV. Light verb constructions in V1	**rakka si yamu* cf. *huri yamu* [fall-do]-stop rain-stop 'stop falling'	[*rakka si*] *owaru* [fall-do]-stop 'stop falling'
V. Idioms in V1	#*abura o uri siburu* oil ACC sell-hesitate 'hesitate to sell oil/ *shoot the breeze'	[*abura o uri*] *sokonau* [oil ACC sell]-fail 'fail to sell oil/shoot the breeze'

Lexical compound verbs consistently exclude syntactic elements from their V1 position whereas syntactic compounds systematically accommodate them. The incompatibility of the former with the syntactically motivated elements endorses the validity of the term "lexical" (or "morphological"), and the compatibility of the latter with them lends credence to labeling it "syntactic". The asymmetries illustrated in Table 1 are so systematic and so robust that the lexical and syntactic groups can be postulated as autonomous classes. This said, two caveats must be noted.

First, it is not the case that all these tests apply blindly to every instance of syntactic compound verbs; rather, their actual applicability is conditioned by such

factors as the semantics of individual verbs and their syntactic structures (Section 3). For example, the verbal anaphora *soo su-* 'do so' in Row III, unlike the English *do so*, is restricted to volitional actions controlled by an agent subject (Matsumoto 1996). In this regard, the embedding of light verbs (Row IV) may be considered the most reliable test as it is exempt from semantic conditioning.

Second, the taxonomy is not intended to assign each compound verb uniquely to either the lexical or the syntactic class; instead, there are several head verbs that have a dual membership. A well-known example is *-kakeru* 'hang' in *hanasi-kakeru* [speak-hang], which means either 'talk to' (lexical) or 'be about to talk' (syntactic). The semantic ambiguity is correlated with the difference in syntactic behavior: only in the latter 'be about to' meaning are the syntactic elements in Table 1 eligible to occur in the V1 position. Likewise, *oki-wasureru* [put-forget] is subject to two different interpretations, depending on whether it is a lexical compound verb ('leave something somewhere and forget to bring it') or a syntactic compound verb ('forget to put something on a designated place'). Only the latter interpretation allows syntactic elements to occur in the V1 position. Other ambiguous head verbs are *-nokosu* 'leave behind' as in *ii-nokosu* [say-leave], which behaves as a lexical compound verb when it means 'leave a remark' and as a syntactic compound verb when it means 'leave unsaid'; *-tukeru* 'attach' as in *nori-tukeru* [ride-attach], which is ambiguous between 'drive up to' (lexical) and 'be used to driving' (syntactic); *-tukusu* 'exhaust', which has a telic interpretation designating completion of an event when used in a syntactic compound, as in *mati o yaki-tukusu* 'burn down all the town', but conveys an atelic, durative meaning when used in the lexical compound verb *tati-tukusu* 'stand still (absent-mindedly)'. The meanings of the dual-membership verbs are neatly correlated with their morphosyntactic behavior.

Now, the syntactic nature of examples like those in (10) above can be further substantiated by such phenomena as scope relations. As observed by Yumoto (2001), *V-sugiru* [V-exceed] 'V excessively', for example, exhibits a range of semantic interpretations that would not obtain, were this compound verb treated as a single word in syntax.

(11) a. *Watasi   wa   wain   o   nomi-sugi-ta.*   (transitive V1)
          I   TOP   wine   ACC   drink-exceed-PST
          'I drank too much wine.'

      b. *Saikin   zisin   ga   okori-sugi-ru.*   (unaccusative V1)
          recently   earthquake   NOM   occur-exceed-NONPST
          'Too many earthquakes have occurred recently.'

      c. *Kodomotati   ga   asobi-sugi-ta.*   (unergative V1)
          children   NOM   play-exceed-PST
          'The children played too much.' NOT 'Too many children played.'

(11a) and (11b) share the interpretation in which -*sugiru* 'exceed' quantifies over the "internal arguments" (i.e. direct objects of transitive verbs and subjects of unaccusative verbs), thus expressing the excessive amount of the wine drunk in (11a) and the excessive number of the earthquakes in (11b). These interpretations are possible because the internal arguments are located inside the scope of the V2 *sugiru* in syntactic structure, as in *[zisin ga okori]-sugiru* ([earthquake NOM occur]-exceed). By contrast, (11c), with the unergative verb 'play', cannot mean that the children who played were too many, because the subject 'children', an external argument, is located outside the scope of *sugiru*, as in *[Kodomotati ga [[asobi]-sugita]]* ([children NOM [[play]-exceeded]]). Due to this structure, (11c) is construed only as designating an excessive amount of the action of playing.

In addition to the syntactic criteria, the two classes of compound verbs differ conspicuously in productivity, lexical restrictions, semantic transparency, and mutual ordering. As regards productivity, lexical compounds are by and large limited to lexically specified combinations of two verbs in V1 and V2 and are thus listed as a whole in dictionaries, although creation of new compounds is by no means precluded. By sharp contrast, syntactic compound verbs are exempt from combinatory restrictions of a lexical nature and hence have no dictionary entries; the only restriction on syntactic compounds is that the heads in V2 are limited to a designated set of 30 auxiliary verbs (Section 3). The difference in productivity was empirically demonstrated by the quantitative study of Tamaoka, Lim, and Sakai (2004), who, applying C.E. Shannon's theory of entropy and redundancy to their analysis of corpus data, found that syntactic compound verbs are significantly higher in entropy than lexical ones though they do not differ markedly in redundancy.

Another notable difference between the two types of compound verbs concerns conditions on vocabulary strata. While lexical compound verbs are strictly restricted to native verbs, syntactic compound verbs have a way to relax the stratum condition, thanks to the light verb constructions that may appear in the V1 position (see Row IV in Table 1). The light verb constructions refer to combinations of a VN followed by *si-*, the infinitive of the native verb *suru* 'do', and the VNs may come not only from the native stratum (e.g. *saka-dati* 'handstand') but more generally from the Sino-Japanese (e.g. *rakka* 'fall') and foreign (e.g. *meeru* 'mail') strata.
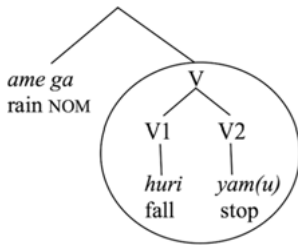
A strict ordering restriction also attests to the validity of the lexical-syntactic distinction. As shown by (12a, b), syntactic compound verbs may follow (i.e. show up outside) lexical compound verbs, but not vice versa.

(12) a. [[lexical]-syntactic] b. *[[syntactic]-lexical]
          [[*huri-yami*]-*kakeru*]   *[[*huri-kake*]-*yamu*]
   [[rain-stop]-be.about.to]   [[rain-be.about.to]-stop]

When first discovered by Kageyama (1989) in the heyday of the controversy over lexicalism and syntacticism, the distinction of the two classes of compound verbs

was ascribed to the demarcation of grammatical modules in which a given compound verb is created. An alternative way of expressing the same dichotomy in a theory-neutral manner is to say that that lexical compounds are lexical (or morphological) because a lexical verb is combined directly with another lexical verb, as shown schematically in (13a), whereas syntactic compounds are syntactic because the head verbs select a syntactic complement clause, as represented in (13b).

(13)   a.   lexical V-V compound   b.   syntactic V-V compound
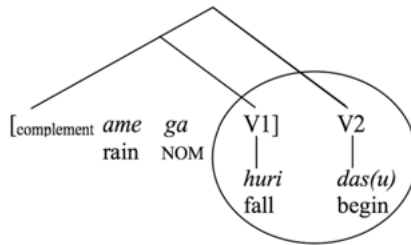
(13a) shows that the two verbs in the circle are directly combined, resulting in a straightforward instance of morphological compounding. (13b), on the other hand, represents a syntactic complementation structure from which the verb in V1 in the complement clause is later integrated with the head verb in V2 by certain mechanisms such as syntactic incorporation (Kageyama 1993), restructuring (Fukuda 2012), or reanalysis (Kishimoto 2013).

Before closing this section, it is intriguing to note that the same morphosyntactic criteria serve to discriminate two classes of Verb-Adjective compound adjectives, as shown in (14).

(14)   a.   Lexical V-A compound adjectives
             musi-atui [steam-hot] 'sultry', *koge-kusai* [burn-smelly] 'smelly of something burning', *ne-gurusii* [sleep-hard] 'not able to sleep well (because of heat)', *mawari-kudoi* [go.round-tedious] 'roundabout'

       b.   Syntactic V-A compound adjectives
             kaeri-tai [go.home-eager] 'eager to go home', *mitome-gatai* [admit-hard] 'hard to accept', *tukai-yasui* [use-easy] 'easy to use', *oboe-nikui* [memorize-hard] 'hard to memorize'

Having the same structure of an infinitive verb followed by an adjective, the two classes of compound adjectives in (14a) and (14b) are indistinguishable morphologically as well as phonologically. Their morphological affinity is reflected in the fact that they are both susceptible to the same derivational expansion with *-sa* (nominalizing suffix) and *-garu* 'show (an internal feeling) outwardly', as in *musi-atu-sa* 'sultriness' (14a-type) and *kaeri-ta-sa* 'eagerness to go home' (14b-type), or *ne-gurusi-garu* 'seem not to sleep well' (14a-type) and *kaeri-ta-garu* 'seem to want to go

home' (14b-type). However, they display distinct behavior when the verb member is replaced by syntactic elements. Table 2 shows that only the compound adjectives of type 14b are compatible with these syntactic expressions in the V position.

Table 2: Lexical vs. syntactic compound adjectives

	Lexical compound adjectives (14a)	Syntactic compound adjectives (14b)
I. Passive in V1	*[*mus are*]-*atui* [steam-PASS]-hot lit. 'hot because of being steamed'	[*ais are*] *tai* [love-PASS]-eager 'eager to be loved'
II. Honorific in V1	*[*o ne ni nari*] *gurusii* [HON-sleep DAT beecome]-hard 'not able to sleep well'	[*o tukai ni nari*] *yasui* [HON-use DAT become]-easy 'easy to use'
III. Verbal anaphora in V1	*[*soo si*] *gurusii* [so do]-hard 'not able to do so'	[*soo si*] *tai* [so do]-eager 'eager to do so'
IV. Light verb constructions in V1	*[*zyukusui si*] *gurusii* [sleep.well-do]-hard 'not able to sleep well'	[*yoonin si*] *gatai* [admit-do]-hard 'hard to admit'
V. Idiom in V1	DNA	[*abura o uri*] *nikui* [oil ACC sell]-hard 'not easily sell oil/shoot the breeze'

As in V-V compound verbs, the syntactic differences between the two types of V-A compound adjectives are correlated with the difference in productivity (lexical compound adjectives are extremely limited in number whereas syntactic compound adjectives are fully productive although the head adjectives are lexically specified to a small number), lexicalization (lexical compound adjectives are registered in dictionaries), and semantic transparency (lexical compound adjectives have fixed meanings).

3 Structure and meaning of syntactic compound verbs

This section outlines the syntactic and semantic properties of syntactic compound verbs by first introducing a semantic classification of their head verbs (Section 3.1) and then presenting their syntactic structures on the basis of their syntactic behavior (Section 3.2).

3.1 Semantic classification of syntactic compound verbs

According to the criteria stated in Section 2.2, thirty verbs are identified as the heads (traditionally called "auxiliary verbs") of syntactic compound verbs (Kageyama 1993; Himeno 1999). Table 3 below classifies them in semantic groups according to the kind of aspectual meaning each head verb conveys; the first verb is indicated simply as "V" because any verb can fit in as long as as the combination is semantically congruent. All the entries are accompanied by their literal meanings in parentheses, followed by their aspectual meanings in single quotes. Comparison of the two meanings will help the reader grasp how the head verbs have undergone semantic bleaching or metaphoric extension. It should be noted, however, that the grammaticalization of the head verbs in V2 stops at the stage of desemanticization (semantic shift to aspectual meanings) and has not developed into decategorization (in the sense of Heine and Kuteva 2002), because these "auxiliary verbs" display the same inflections as full-fledged verbs.

Table 3: Syntactic compound verbs classified semantically

Aspectual notions	Examples
inception	V-*kakeru* [hang, vt.] 'be about to V', V-*dasu* [take out, vt.] 'begin to V', V-*hazimeru* [begin, vt.] 'begin to V', V-*kakaru* [set about, vi.] 'set about V-ing'
duration, repetition	V-*tuzukeru* [continue, vt.] 'continue to V', V-*makuru* [roll up, vt.] 'V on and on'
completion	V-*oeru* [finish, vt.] 'finish V-ing', V-*owaru* [end, vi.] 'stop V-ing', V-*tukusu* [exhaust, vt.] 'V exhaustively', V-*kiru* [cut, vt.] 'V completely', V-*toosu* [let through, vt.] 'V to the end', V-*nuku* [pull out, vt.] 'V to the end', V-*hateru* [come to an end, vi.] 'V utterly'
incompletion, failure	V-*sokonau* [harm, vt.] 'miss V-ing', V-*sokoneru* [harm, vt.] 'fail to V', V-*sonziru* [damage, vt.] 'fail to V', V-*sobireru* [miss (auxiliary use only)] 'fail to V', V-*kaneru* [be.unable] 'be unable to V', V-*wasureru* [forget] 'forget to V', V-*nokosu* [leave undone] 'leave something without V-ing completely', V-*ayamaru* [err, vt.] 'make a mistake in V-ing', V-*okureru* [be late, vi.] 'be delayed in V-ing', V-*aguneru* [Classical Japanese 'be satiated'] 'hesitate to V'
excessiveness	V-*sugiru* [go past, vi.] 'V excessively'
retrial	V-*naosu* [repair, vt.] 'V again (to obtain a desired result)'
repetition, habituation	V-*tukeru* [attach, vt.] 'be used to V-ing', V-*nareru* [get accustomed, vi.] 'be accustomed to V-ing', V-*akiru* [get weary, vi.] 'get weary of V-ing'
reciprocity	V-*au* [meet, vi.] 'V reciprocally'
likelihood	V-*eru* or V-*uru* [obtain, be possible] 'be likely to V'

While Matsumoto (1996) and Fukuda (2012) restrict the range of "aspectual verbs" narrowly to *-hazimeru* 'begin', *-owaru* 'end', *-oeru* 'finish', *-kakeru* 'be about to', and a couple others, we take it that the head verbs in Table 3 converge on the broad concept of lexical aspect or *Aktionsarten* (lit. 'manner of action') which covers a multiplicity of notions related to how an event does or does not unfold. In Section 4, an even wider range of *Aktionsart* meanings will be attested in lexical compound verbs.

3.2 Structures of syntactic compound verbs

Japanese has a huge spectrum of syntactic suffixes expressing a diversity of meanings that would be represented by independent verbs, auxiliaries, or prefixes in other languages, such as causative *-sase(ru)* 'cause', direct and indirect passive *-rare(ru)*, desiderative *-ta(i)* 'want', actional *-gar(u)* 'exhibit one's feeling by action', and ease/difficulty *-yasu(i)* 'easy'/*-niku(i)* 'hard'. These suffixes interact not only with each other but also with syntactic compound verbs. Since the inception of Japanese generative grammar, it has been almost universally assumed that such syntactic suffixes are independent predicates (verbs or adjectives) that take a complement clause denoting an event or state (Kuroda 1979; Kuno 1973; Inoue 1976; Shibatani 1978; Sugioka 1986; Miyagawa 1989; among others). In addition to a complement clause, many syntactic suffixes require their own subject, typically a human agent or experiencer. The causative *-sase(ru)* and actional *-gar(u)*, for example, select a volitional agent on their own, whereas the desiderative *-ta(i)* calls for its own experiencer subject, as in [*Watasi wa [~~watasi ga~~ kaeri]-ta-i]* (I want [~~I~~ go home]) 'I want to go home.' This structure, corresponding to the English construction *Mary is eager to go*, is called "control structure". On the other hand, *-sooda* 'be likely', as in *Ken ga ki-sooda* 'Ken is likely to come' and *Ame ga huri-sooda* 'It is going to rain', is not limited to a volitional human subject, thus motivating "raising structure" like *[Ame ga huri]-sooda.* ([Rain falls] is likely), on a par with the English sentence *It is likely to rain*.

This structural distinction can be applied to syntactic compound verbs as well.

(15)   a.   Raising (or intransitive) structure
         [*Zisin*   *ga*   *okori*]-*kake-ta.*
         [earthquake   NOM   happen]-be.about.to-PST
         'An earthquake was about to occur.'

       b.   Control (or transitive) structure
         Haha   *wa*   [~~*haha*~~   ~~*ga*~~   *asagohan*   *o*   *tabe*]-*sobire-ta.*
         mother   TOP   [~~mother~~   ~~NOM~~   breakfast   ACC   eat]-miss-PST
         'Mother missed eating breakfast.'

Example (15a) is analyzed as a raising (or intransitive) structure because its head verb *-kake(ru)* 'be about to' is capable of predicating either a natural event with an inanimate subject or a volitional action with a human subject. By contrast, (15b) is given a control (or transitive) structure because the head verb *sobire(ru)* 'miss' strictly requires a volitional human subject.

Drawing on the distinction between raising and control structures, Kageyama (1993) differentiated three classes of head verbs in V2 as regards the type of complement clause each V2 selects. The three classes are summarized in (16) together with their argument structures and examples.

(16)　a.　VP complement in raising structure
　　　　－ V2's argument structure: <Proposition>
　　　　－ Examples: V-*kakeru* 'be about to V', V-*dasu* 'start to V', V-*sugiru* 'V excessively', V-*eru/uru* 'be likely to V', V-*owaru* 'finish V-ing', V-*kanenai* 'something bad is likely to happen'

　　　b.　VP complement in control structure
　　　　－ V2's argument structure: (Experiencer/Agent <Proposition>)
　　　　－ Examples: V-*sobireru* 'miss V-ing', V-*tukeru* 'be used to V-ing', V-*kaneru* 'be reluctant to V', V-*ayamaru* 'make an error in V-ing', V-*aguneru* 'hesitate to V', V-*nareru* 'be accustomed to V-ing', V-*akiru* 'get weary of V-ing', V-*toosu* 'V to the end'

　　　c.　V' (reduced) complement
　　　　－ V2's argument structure: (Agent/Experiencer <Theme>)
　　　　－ Examples: V-*oeru* 'finish V-ing', V-*naosu* 'V again', V-*tukusu* 'V exhaustively', V-*sonziru* 'fail to V', V-*sokonau* 'fail to V', V-*wasureru* 'forget to V', V-*nokosu* 'V and leave half-done', V-*nuku* 'V to the end', V-*au* 'V mutually'

　　　d.　Dual membership of 16a-type structure and 16c-type structure
　　　　V-*hazimeru* 'begin to V', V-*makuru* 'V on and on', V-*tuzukeru* 'continue to V', V-*kiru* 'V completely'

In Kageyama (1993, 1999), the three structures of (16a), (16b), and (16c) were analyzed as schematically represented in Figures 1, 2, and 3, respectively.

These tree diagrams, based on the VP-internal subject hypothesis (subject is projected directly under VP, and object under V´) in the generative grammar of the early 1990s, would be easily translated in the split verb phrase structure in current Minimalist syntax by replacing the categories VP and V´ with "vP" and "VP", respectively (cf. Fukuda 2012).

Let us begin with the difference between the raising structure in Fig. 1 and the control structure in Fig. 2. The head verb (V2) in the former does not impose any
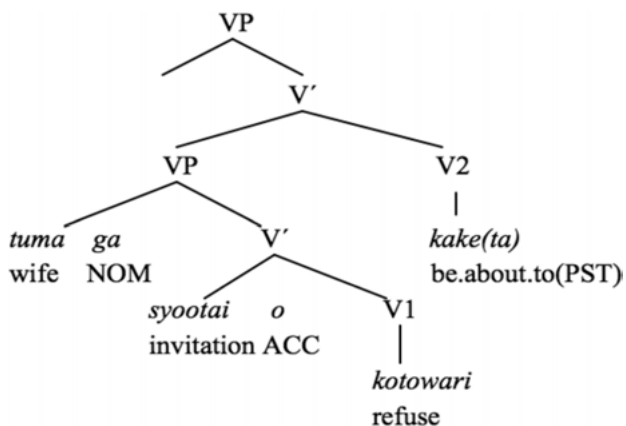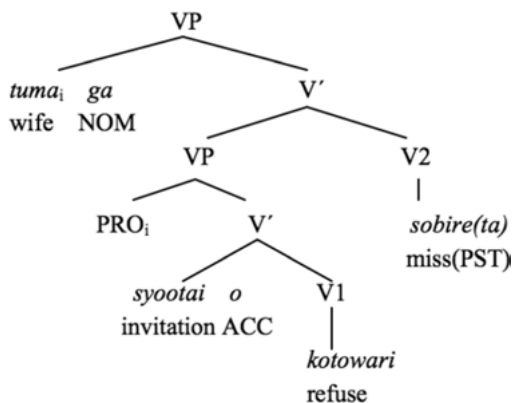
Figure 1: Raising structure

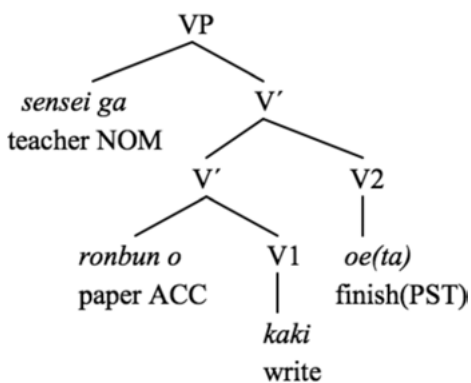**Figure 2:** Control structure

Figure 3: Reduced complement structure

semantic restriction on the subject, as shown in (17), while that in the latter requires the subject to be a sentient human, as shown in (18).

(17) a. *Zisin   ga   okori-kake-ta.*   (inanimate subject)
          earthquake   NOM   occur-be.about.to-PST
          'An earthquake was about to occur.'

    b. *Tuma   ga   syootai   o   kotowari-kake-ta.*   (human subject)
          wife   NOM   invitation   ACC   refuse-be.about.to-PST
          'My wife was about to turn down the invitation.'

(18) a. **Zisin   ga   okori-sobire-ta.*   (inanimate subject)
          earthquake   NOM   occur-miss-PST
          lit. 'An earthquake missed the chance of occurring.'

    b. *Tuma   wa   syootai   o   kotowari-sobire-ta.*   (human subject)
          wife   TOP   invitation   ACC   reject-miss-PST
          'My wife missed the chance of turning down the invitation.'

Both raising and control structures have a bi-clausal character because the main and complement clauses are separated by VP. In neither structures does the head verb in V2 have the right to impose any selectional restriction on the embedded object. Instead, the object argument is associated exclusively with the verb in V1. Consequently, the verb in V2 can never be passivized, although that in V1, if transitive, can. This point is demonstrated by the examples in (19) using a raising-type V2 *-kakeru* 'be about to' and by those in (20) using a control-type V2 *-nareru* 'be used to'.

(19) a. *Watasi   wa   siranai   otoko   ni   damas-**are**-kake-ta.*   (passive on V1)
          I   TOP   strange   man   by   deceive-PASS-be.about.to-PST
          'I was about to be deceived by a stranger.'

    b. **Watasi   wa   siranai   otoko   ni   damasi-kake-**rare**-ta.*   (passive on V2)
          I   TOP   strange   man   DAT   deceive-be.about.to-PASS-PST

(20) a. *Watasi   wa   titi   ni   sikar-**are**-nare-te   i-ru.*   (passive on V1)
          I   TOP   father   by   scold-PASS-be.used-GER   be-PRS
          'I am used to being scolded by my father.'

    b. **Watasi   wa   titi   ni   sikari-nare-**rare**-te   i-ru.*   (passive on V2)
          I   TOP   father   by   scold-be.used-PASS-GER   be-PRS

The ungrammaticality of (19b) and (20b) is a natural consequence of their "bi-clausal" structures.

With respect to the selection of subjects, the reduced-complement type in (16c), shown in Fig. 3, exhibits the same pattern as the control-type (16b), thus requiring an agent subject on its own (Shibatani 1973; Kuno 1983; Nishigauchi 1993; Kageyama 1993; Matsumoto 1996; Koizumi 1999). Fukuda (2012) questions this view by adducing examples with non-human subjects like (21), but this is not a good argument.

(21) a. *Densya ga   hasi   o   watari-oe-ta.*
          train   NOM   bridge ACC   cross-finish-PST
          'The train finished crossing the bridge.'

   b. *Siboo o   tokasi-oe-ta   yakuhin   wa* ...
       fat   ACC   melt-finish-PST   chemical TOP ...
       'the chemical that finished melting the fact ...'

Human-controlled artifacts like 'trains' and 'chemicals' should be treated on a par with their manipulators (volitional human agents) because of their metonymic relationship.

Another disputable case is V-*owaru* 'finish' (vi.), classified in the raising type in (16a). There are cases where this verb is incompatible with inanimate subjects, as in (22a).

(22) a. **Ame   ga   huri-owat-ta.*
          rain   NOM   fall-finish-PST
          lit. 'Rain finished falling.'

   b. *Suna-dokei no   suna ga   oti-owat-ta.*
       sand-clock   GEN   sand NOM   fall-finish-PST
       'The sand of the sandglass finished falling.'

The ungrammaticality of (22a), however, should be contrasted with the increased acceptability of (22b), an actually occurring example. The contrast between 'rain falling' and 'sand falling' appears to tie in with Pustejovsky's (1995: 206) observation that the acceptability of the English verb *finish* differs in **It has finished raining* and *The leaves have finished falling*. Pustejovsky's solution was to regard the leaves, but not rain, as a "pseudo-agent", arguing that such "pseudo-control" cases are restricted to a certain class of nouns. An alternative way of explaining the contrast would be to maintain that *owaru* (or *finish*) takes a raising structure syntactically but it imposes a semantic restriction on the aspectual property of the complement clause in such a way that the event in V1 must have a predetermined endpoint. On this account, (22a) is ruled out because one cannot tell beforehand when the rain will stop, whereas (22b) makes sense because the end of the sand's dropping is predetermined by the capacity of the sandglass. In fact, sentences like *Tanku no mizu ga de-owatta* [tank GEN water NOM flow.out-finished] 'The water in the tank

finished flowing out' sound natural because it is easily expected that the flowing of the water will end at the time when the tank becomes empty. The requirement for a predetermined endpoint, which is absent from verbs of inception like V-*dasu* 'start' and verbs of continuation like V-*tuzukeru* 'continue', appears to be a special condition on verbs of completion, including V-*kiru* 'V completely' (16c). Note that the lexical compound *huri-yamu* 'stop raining' is immune from this condition.

Now, the head verbs of the reduced complement-type (16c) display an amazing property of inducing what Nishigauchi (1993) called "long distance passive". By applying to the verb in V2 rather than in V1, long distance passive promotes the object in the embedded clause to the matrix subject. While Nishigauchi (1993) focused on -*oeru* 'finish', Kageyama (1993) extended his observation to all the head verbs of this type and proposed the reduced V´-complement structure in Fig. 3, where V1 projects only to a V´ structure that lacks its own subject. The syntactic structure of this type of compound verb, which has only one subject, is mono-clausal. Because of this, the embedded object can be passivized at the main clause (following the Relativized Minimality of the Government and Binding theory). The examples in (23) illustrate how long distance passive applies to compound verbs of the mono-clausal type.

(23) a. *Koozan ga   hori-tukus-**are**-ta.*
      mine   NOM dig-exhaust-PASS-PST
      'The mine was dug up and exhausted.'

   b. *Keiyakusyo ga   kaki-naos-**are**-ta.*
      contract   NOM write-do.again-PASS-PST
      'The contract was rewritten and fixed.'

   c. *Dasi-wasure-**rare**-ta tegami ga   tukue no   ue ni   ar-u.*
      post-forget-PASS-PST letter   NOM desk GEN top DAT be-NONPST
      'There is a letter on the desk that he forgot to post and left unwittingly.'

Kageyama (1993) proposed to attribute the possibility of long distance passive to the lexical property of the head verbs themselves. Although these V2s have acquired aspectual notions, they still appear to retain their original lexical meanings at the same time, as summarized in Table 4.

Table 4: Literal and aspectual meanings

	literal meaning	derived aspectual meaning
-*oeru*	'finish, complete'	'finish doing, so that the object is done'
-*naosu*	'fix, repair'	'do again, so that the object is in a fixed state'
-*wasureru*	'forget'	'forget to do, so that the object is left forgotten'
-*tukusu*	'exhaust'	'do thoroughly, so that the object is exhausted'

The literal and derived meanings summarized in Table 4 suggest that these head verbs play a dual role of behaving as an aspectual verb taking V1's action in scope and as a lexical verb that assigns a particular meaning to the object. For example, *syoosetu o kaki-oeru* [novel ACC write-finish] is interpreted as 'write a novel and finish it' or 'finish a novel by writing'. The thematic contribution of the second verb becomes more apparent in the paired examples of (24), taken from Fukui, Miyagawa and Tenny (1985), where the semantic scope of V2 *tukusu* 'exhaust' varies according to what comes in the object position.

(24) a. *Masao wa kabe ni posutaa o hari-tukusi-ta.*
         Masao TOP wall DAT poster ACC aste-exhaust-PST
         'Masao pasted all the posters onto the wall.'

      b. *Masao wa posutaa de kabe o hari-tukusi-ta.*
         Masao TOP poster with wall ACC paste-exhaust-PST
         'Masao pasted the entire wall with posters.'

In (24a), the posters were exhausted (used up), and in (24b), the surface of the wall was covered up.

Considerations like these led Kageyama (1993) to postulate that in the syntactic structure of Fig. 3, the two verbs in V1 and V2 jointly determine the semantics of the object in the manner of Baker's (1989) double theta-role assignment (see also Kishimoto (2013) for an alternative analysis of long distance passive).

Returning to the taxonomy in (16), the verbs shown in (16d) have a dual membership of raising type (16a) and reduced complement type (16c), and their ability to passivize varies accordingly, as observed by Matsumoto (1996) and Fukuda (2012).

(25) a. *Seisyo wa eien ni yom-**are**-tuzuke-ru daroo.* (Type 16a)
        bible TOP eternity DAT read-PASS-continue-PRS will
        'The Bible will continue to be read eternally.'

      b. *Seisyo wa eien ni yomi-tuzuke-**rare**-ru daroo.* (Type 16c)
        bible TOP eternity DAT read-continue-PASS-PRS will
        lit. 'The Bible will be continued to read eternally.'

The sentence in (25a), associated with the raising structure of (16a), means that the event of the Bible being read will continue forever, where the agent's volitional control is limited to the reading act denoted by V1. On the other hand, the sentence in (25b), reflecting the mono-clausal structure of (16c), conveys that the Bible has an eternally continued readership, where the agent's volitional control extends to V2's event of continuing the reading action. This semantic difference between passive on V1 and passive on V2 has its syntactic reflections in (26a) and (26b).

(26) a. ?*_Seisyo   wa   Taroo   ni(yotte)   yom-are-tuzuke-ta._
       bible   TOP   Taro   by   read-PASS-continue-PST
       'The Bible continued to be read by Taro.'

    b. _Seisyo   wa   Taroo   ni(yotte)   yomi-tuzuke-rare-ta._
       bible   TOP   Taro   by   read-continue-PASS-PST
       lit. 'The Bible was continued by Taro to read.'
       = 'Taro intentionally continued reading the Bible (while no one else did so).'

As pointed out by Masuoka (1987), the agents of the passive _yom-are_ 'be read' may not be a specific individual like Taro, as shown by the unnaturalness of ?*_Seisyo wa kinoo Taroo ni(yotte) yom-are-ta_ 'The Bible was read by Taro yesterday'. This unnaturalness is carried over to the low acceptability of 'by Taro' in (26a), where V1 'read' is passivized. In contrast, the agent phrase 'by Taro' in (26b), where passive applies to V2 'continue' (or possibly to the whole compound V1-V2), gains increased acceptability because Taro is interpreted as controlling the event of V2 _tuzukeru_ 'continue', whose passive form is in fact compatible with a specific agent, as in _Sono sigoto wa Taroo niyotte tuzuke-rare-ta_ 'The work was continued (taken over) by Taro'. These observations indicate that the semantic property of V1 plays a vital role in the "downstairs" passive, and that of V2 in the "upstairs" passive.

    While it is generally agreed that a three-way distinction like the one presented in (16a, b, c) is empirically cogent, opinions diverge as to exactly how it is implemented in theoretical frameworks. In the framework of Lexical-Functional Grammar, for example, Matsumoto (1996) argues that the raising and control types with a VP complement are bi-clausal in both functional structure and argument structure while the V´-complement type is mono-clausal in functional structure though bi-clausal in argument structure. Yumoto (2005, 2009), on the other hand, entertains the idea that the two members of mono-clausal type compound verbs directly merge in syntax as a compound verb of the form $[V^0 \ V^0]_V$. Such a structure, however, will have difficulty in explaining the occurrence in the V1 position of syntactic elements such as causative, honorific, light verbs, and idioms. Thus Yumoto's analysis is essentially the same as treating the syntactic compounds of the mono-clausal type as lexical and is therefore incapable of explaining the crucial differences between syntactic and lexical compounds.

    Subject honorification also serves to differentiate the three structural types. Shibatani (1978) and Kuno (1983) pointed out that the aspectual verb _hazimeru_ 'begin' allows the honorific marker _o…ni naru_ to be attached either to V1 alone or to the whole V1-V2.

(27) a. _Sensei   wa   tegami   o   o-[_{V1} kaki]   ni   nari   [_{V2} hazime-ta]._
       teacher   TOP   letter   ACC   HON-write   DAT   become   begin-PST
       'The teacher began to write a letter.'

    b. _Sensei   wa   tegami   o   o-[_{V1} kaki]-[_{V2} hazime]   ni   nat-ta._
       teacher   TOP   letter   ACC   HON-write-begin   DAT   become-PST

Shibatani and Kuno attributed (27a) and (27b) to raising and control structures, respectively. Although (27a) is unequivocally a case of raising structure, (27b) should be identified as the reduced complement type because the subject 'teacher' is associated with both V1 'write' and V2 'begin'. In fact, the raising verbs of (16a) generally reject the honorification pattern of (27b). Examples like the one in (28) are unacceptable to many speakers, although the judgment may vary depending on individual speakers (Kuno 1983).

(28) *Sensei   wa   tegami o   o-[kaki-owari]   ni   nat-ta.   (Matsumoto 1996)
     teacher TOP letter   ACC HON-write-finish DAT become-PST
     'The teacher finished writing a letter.'

On the other hand, the compound verbs with mono-clausal structure largely follow the honorification pattern of (27b), as in o-[kaki-wasure] ni naru 'forget to write' and o-[kaki-naosi] ni naru 'rewrite'. Note incidentally that the head verbs of this type also allow the honorific marker to be attached only to V1, at least for some speakers (Kageyama 1993).

(29) ?Sensei wa   denwa o   o-kake   ni   nari   naosi-ta/oe-ta.
     teacher TOP phone   ACC HON-make DAT become do.again-PST/finish-PST.
     'The teacher made a phone call again/finishing speaking on the phone.'

The relatively high acceptability of (29) runs counter to Yumoto's (2009) base generation analysis of this type of compound verb.

    To summarize, theoretical research on the structure of syntactic compound verbs based on the lexical-syntactic distinction initiated by Kageyama (1993) was followed up by Yumoto (2001, 2005, 2009), who discovered a number of interesting semantic (mostly, scope-related) phenomena that reinforce the structural distinctions among the three groups of syntactic compound verbs. On the other hand, Matsumoto (1996), couched in the framework of Lexical-Functional Grammar, elaborated on Kageyama's three-way taxonomy by utilizing argument structure and functional structure as two independent representations. More recently, Fukuda (2012), focusing on four aspectual verbs (hazimeru 'begin', owaru 'finish', oeru 'complete', tuzukeru 'continue'), proposed to treat them as functional heads that project Aspect Phrases in two different layers of syntactic structure, High-aspect and Low-aspect. While Fukuda (2012) brought to light intriguing data involving the compatibility of these aspectual verbs with the event types of their complement clauses, it is not immediately clear whether such semantic observations provide direct evidence for the decategorization of the head verbs to the functional category of "aspectuals" as Fukuda claims.

4 Structure and meaning of lexical V-V compounds

This section outlines the architecture of lexical V-V compound verbs. In 4.1, two types of lexical compound verbs – thematic and aspectual – are distinguished in terms of the functional grammaticalization of the verbs in V2. The characteristics of lexical aspectual compound verbs are delineated in 4.2, and those of thematic compounds verbs in 4.3. In 4.4, the combinatory restrictions of two verbs are discussed chiefly in light of Kageyama's (1993) Transitivity Harmony Principle.

4.1 Distinction between thematic and aspectual lexical compounds

At first glance, lexical compound verbs appear to be a simple concatenation of two verbs. In fact, the core meanings of approximately 60% of about 2,750 lexical compounds in my database (available online as Kageyama and Kanzaki 2014) can be plausibly paraphrased by replacing the infinitive verb in V1 with its gerundive (-*te*) counterpart in the formula [V1-V2] ≑ 'V1-*te* V2' [V1 and then V2], where V1 modifies V2 from left to right (see also Nagashima 1976). For example, *osi-akeru* [push-open (vt.)] is virtually synonymous with *osi-te akeru* 'open by pushing, push open', *nage-ireru* [throw-put.in] with *nage-te ireru* 'put in by throwing, throw in', and *asobi-kurasu* [play-live] with *ason-de kurasu* 'spend a day playing, idle one's time away'. The paraphrase by 'V1-*te* V2' is reasonable because it is isomorphic with the temporal progress of the two events in V1 and V2.

Curiously, however, approximately 30% of the lexical compound verbs available in my database resist such a paraphrase. On the contrary, many of them prefer a paraphrase with a reversed order in which V2 modifies V1 from right to left. For example, *huri-sikiru* [fall-continue] '(rain or snow) fall incessantly' cannot be paraphrased as **hut-te sikiru* (in fact, the verb *sikiru* is limited to the V2 position of a few lexical compound verbs in contemporary Japanese) but instead as *sikiri ni huru* [do.repeatedly DAT fall] 'fall incessantly' by reversing the order of the two members. Likewise, *(hana ga) saki-kisou* [bloom-compete] '(flowers) be in full blossom' can be paraphrased most plausibly as *kisot-te saku* 'bloom competitively', and *hare-wataru* [be.cloudless-go.over] '(sky) be clear all over' as *sumizumi ni watat-te hareru* [throughout.the.sky DAT go.over be.cloudless]. There are also verbs that fit in neither with the 'V1-*te* V2' formula nor with the reversed order paraphrase, such as *saki-someru* [bloom-begin] 'begin to bloom' and *tukuri-ageru* [make-complete] 'finish making'. In these compounds, the verbs in V2 add various *Aktionsart* meanings to the verbs in V1.

The disparity in paraphrasability points to the need to divide what have been subsumed under the rubric of lexical compound verbs into two classes. Specifically, the two classes differ in the way the argument structure of a whole compound is

determined. Let us use the term "thematic compound verbs" to refer to the group in which the paraphrase with 'V1-*te* V2' is semantically appropriate, and "aspectual compound verbs" to refer to the group where the same formula does not apply (Kageyama 2013). In the thematic group, each of the two members is equipped with its own argument structure showing the thematic relations of subject and object. In *osi-akeru* [push-open], for example, each of V1 *osi-* 'push' and V2 *akeru* 'open (tr.)' has an argument structure, and V1 semantically modifies the head V2. Since these two verbs are both transitive, they share the argument structure <agent, theme>. The agents of V1 and V2 are identified as referring to the same individual, and the themes of the two verbs are likewise identified as referring to the same entity. In Section 4.3, we will explain in some detail how the arguments of the two members are unified and how the selection restrictions of the two verbs are integrated.

The class of lexical aspectual compound verbs exhibits a striking contrast to the thematic class in many respects, most importantly in argument realization. The fact that aspectual compound verbs refuse a paraphrase by the formula 'V1-*te* V2' indicates that V2 is basically devoid of argument structure; rather, the argument relation of a whole compound pivots on the verb in V1, with V2 supplying it with semantic embellishment. In *hare-wataru* [become.clear-go.over] '(sky) be clear all over', for example, the argument relation is regulated by V1 *hare-* 'become cloudless', to which V2 *wataru* adds an aspectual meaning of the completeness of the event in V1. Likewise, in *(hana ga) saki-midareru* [flowers NOM bloom-be.confused] '(flowers) bloom in great profusion', the subject 'flowers' is the argument of V1 *saku* 'bloom', and V2 *midareru* lit. 'become confused, become untidy' has lost its literal meaning, conveying only an adverbial meaning like 'in great (lit. confused) profusion' that depicts the way the subevent in V1 unfolds. Such an aspectual meaning is reminiscent of the head verbs of the syntactic compound verbs explained in Section 3. It must be stressed, however, that the aspectual compounds under discussion are formed lexically rather than syntactically because they fail in all the diagnoses for syntactic compounds (Section 2.2).

As in the case of syntactic compound verbs, dual memberships are observed in the domain of lexical compounds, where one and the same verb in V2 makes up a thematic compound on the one hand and an aspectual compound on the other. There are quite a few verbs in the V2 position that exhibit such a thematic/aspectual ambiguity. For example, the verb *-ageru* 'raise, send up', which is fairly productive as a V2 of lexical compounds, expresses different meanings in (30a) and (30b).

(30) a. *Watasi   wa   kaban   o   moti-age-ta.*
          I   TOP   bag   ACC hold-raise-PAST
          'I lifted up the bag.'

      b. *Watasi   wa   ronbun   o   kaki-age-ta.*
          I   TOP   paper   ACC write-finish-PAST
          'I finished writing a paper.'

The compound verb *moti-ageru* 'lift up' in (30a), paraphrased as *mot-te ageru* 'raise by holding', is thematic inasmuch as its two members both represent physical actions of holding a bag and raising it. The compound *kaki-ageru* 'finish writing' in (30b), on the other hand, is aspectual because V1 *kaki-* 'write' determines the argument relation of the whole sentence and V2 *ageru* (lit. 'raise') adds the aspectual meaning of completion to the event in V1.

Another instance of dual membership is the verb *-komu* '(move) in', which is a bound root that turns up only as the second member of lexical V-V compound verbs. The transitivity of a whole compound that contains *-komu* in V2 is determined entirely by the verb in V1. Now, the examples in (31a, b) exemplify the thematic and aspectual usages of *-komu*, respectively.

(31) a. *Yakuza   ga   mise ni   abare-kon-da.*
     hoodlum NOM shop DAT rage-go.in-PST
     'A hoodlum (or hoodlums) stormed into the shop.'

   b. *Taikai   o   mokuzen   ni   hikae-te, sensyu wa   hasiri-kon-da.*
     meeting ACC right.before DAT wait-GER athlete TOP run-do.a.lot-PST
     'With a meeting close at hand, the athlete(s) did a lot of running training.'

The compound verb in (31a) is thematic because V2 *-komu*, albeit a bound root, adds a Goal argument to V1, thereby making the event telic. By contrast, the same compound verb in (31b) is classified as an aspectual compound because *-komu* does not add a goal phrase but instead expresses iteration of a running event, thus rendering the whole event in (31b) atelic.

In addition to the differences in meaning and argument realization, Kageyama (2013) observes a restriction on the ordering of aspectual and thematic compounds within the broad category of lexical compound verbs: An aspectual V2 may appear after a thematic compound, as in (32a), but the reversed order is disallowed, as in (32b).

(32) a. [*ne-sizumari* (THEMATIC)]-*kaeru* (ASPECTUAL)
     [sleep-get.silent]-do.thoroughly
     'become thoroughly silent after people fall asleep'

   b. *[*ne-kaeri* (ASPECTUAL)]-*sizumaru* (THEMATIC)
     [sleep-do.thoroughly]-get.silent

This ordering restriction leads us to hypothesize that an aspectual compound is generated above a thematic compound in a hierarchical morphological structure.

(33)
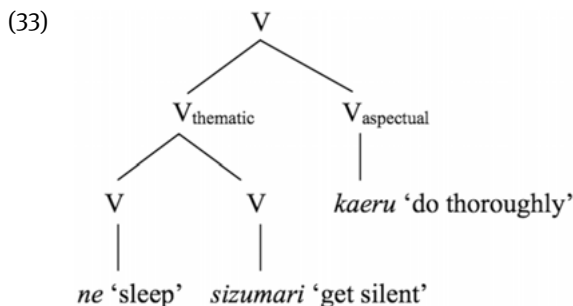
The stacking relation in (33) is in line with the cognitive structure of human language in which an aspectual notion (-*kaeru* 'thoroughly') takes scope over a thematic or propositional content (*ne-sizumaru* 'become silent after people went to sleep').

4.2 Lexical aspectual compound verbs

Martin (1975) provides a list of as many as 56 infinitive-attached auxiliary verbs including nominal and adjectival auxiliaries. A close inspection reveals, however, that his list is a mixture of syntactic and lexical aspectual verbs. For convenience' sake, the V2s of lexical aspectual compound verbs will be referred to as "L-aspectual verbs", as opposed to syntactic aspectual verbs like -*hazimeru* 'begin'.

L-aspectual verbs are characterized by a number of idiosyncratic properties, many of which are shared by the syntactic aspectual verbs.

(34)   L-aspectual verbs are characterized by:

a.   their function to provide V1 with a wide variety of *Aktionsart* meanings such as inception, continuation, completion, cessation, failure, strength, etc.;

b.   their broad semantic coverage of not only temporal aspect but also spatial aspect/deixis, as in *hare-wataru* [be.cloudless-extend.all.over] 'The clear state extends all over the sky', and interpersonal aspect/deixis, as in *moosi-ageru* [say-give.to.a.person.of.higher.status] 'say deferentially';

c.   their peculiar abstract and metaphorical meanings that are not found in their original, independent usage meanings;

d.   the general lack of argument structure and case;

e.   their highly selective cooccurrence relations with particular V1s;

f.   the lack of independent usage in some verbs.

Representative examples of L-aspectual verbs are shown in Table 5 together with their characteristic meanings classified into temporal, spatial, and social aspect. This classification subsumes both those compound verbs that were characterized as complementation-type by Kageyama (1993) and Yumoto (2005) and those that were characterized as "adverbial" by Matsumoto (1998).

Table 5: Lexical aspectual compound verbs classified semantically

I. Temporal aspect a. completive	*huri yamu* (vi.) 'stop falling, said of snow and rain', *kaki ageru* (vt.) 'finish writing', *kaki agaru* (vi) 'writing is finished', *ni tumeru* (vt.) 'boil down', *kesi saru* (vt.) 'wipe out'
b. incompletive	*ii sasu* (vt.) 'stop speaking halfway'
c. intensive result	*ne komu* (vi.) 'fall sound asleep', *komari hateru* (vt.) 'be completely at a loss', *sizumari kaeru* (vi.) 'become completely silent', *saki midareru* (vi.) 'bloom in profusion', *otituki harau* (vi.) 'be perfectly composed'
d. inception	*ake someru* (vi.) 'begin to dawn', *saki someru* (vi.) 'begin to bloom'
e. continuative	*huri sikiru* (vi.) '(rain) fall incessantly', *naki kurasu* (vi.) 'cry all day'
f. iterative	*hozikuri kaesu* (vt.) 'dig again', *tukai komu* (vt.) 'use repeatedly', *tate kaeru* (vt.) 'rebuild', *ii narawasu* (vt.) 'commonly say'
g. intensive action	*sawagi tateru* (vi.) 'fuss about', *izikuri mawasu* (vt.) 'fumble about', *waki kaeru* (vi.) 'boil violently', *waki tatu* (vi.) 'boil hard', *home tigiru* (vt.) 'praise highly'
h. ineffective	*kasi siburu* (vt.) 'hesitate to lend', *nobi nayamu* (vi.) 'do not make expected progress', *kiki tigau* (vt.) 'hear wrongly', *haki tigaeru* (vt.) 'put on the wrong shoes'
i. reciprocal	*i awaseru* (vi.) 'happen to be at the same place', *kiki kaesu* (vt.) 'ask back', *tukai wakeru* (vt.) 'use different things according to the needs'
II. Spatial aspect	*naguri kakaru* (vt.) 'strike at', *donari tukeru* (vt.) 'yell at', *mi ageru* (vt.) 'look up', *mi mawasu* (vt.) 'look round', *hare wataru* (vt.) 'be clear all over the sky', *tobi dasu* (vi.) 'dash out'
III. Social (interpersonal) aspect	*moosi ageru* (vt.) 'say to a respectable person', *mi kudasu* (vt.) 'look down upon', *moosi tukeru* (vt.) 'say to a person of lower status', *kasi sageru* (vt.) '(public institution) lend to private sector'

Table 5 shows that Japanese L-aspectual verbs cover a far wider range of aspectual meanings than the verb-particle combinations in English do. Although the two languages share many aspectual notions such as continuation (*huri-sikiru* 'rain **on and on**'), completion (*kaki-ageru* 'write **up**'), intensive result (*hare-wataru* '(The sky) clear **up**'), and intensive action (*sawagi-tateru* 'fuss **about**'), Japanese compound verbs are unique in expressing such notions as the incompletion, ineffectiveness, or reciprocality of actions, and some L-aspectual verbs have the additional function of inducing lexical valency change (Section 4.4).

In contrast to the diversity of the internal semantic relations between V1s and L-aspectual verbs, their external syntactic properties are defined in a rather simple manner. Since L-aspectual verbs lack a full-fledge argument structure, the argument relation and transitivity of the whole aspectual compound are determined by the verbs in V1, regardless of the morphological transitivity of V2s. Thus *tobi-dasu* [jump (vi.)-take.out (vt.)] 'jump out' is intransitive despite the transitive morphology of the L-aspectual *-dasu* lit. 'take out', and so is *i-awaseru* [be (vi.)-put.together (vt.)] '(two or more people) happen to be at the same place'.

The semantic interpretation of lexical aspectual compound verbs will proceed as illustrated in (35) with the L-aspectual verb *sikiru* 'take place incessantly', a bound morpheme that appears only in aspectual compound verbs.

(35) a. Lexical representation of *sikiru*
          i. subcategorization: [*huri* 'fall' / *naki* 'chirp']
          ii. semantics: INCESSANTLY(e)

      b. Semantics of compound verb *huri-sikiru*
          huri '(rain/snow) fall': **fall´** (e, RAIN/SNOW)
          huri-sikiru: **fall´** (e, RAIN/SNOW) & INCESSANTLY(e)
          = INCESSANTLY(**fall´** (e, RAIN/SNOW))
          (The event of rain/snow falling takes place incessantly.)

In (35), the formal semantic formulas are given only for illustrative purposes. What is important is that the semantic representation of V1 ('fall' in the above example) remains intact and is modified by the L-aspectual verb *sikiru*, which functions as an event modifier. Because of this, the argument structure of a whole compound is determined, in principle, by the verb in V1 alone.

Exceptional cases are found where not just the verb in V1 but also the L-aspectual verb appear to make contribution to the determination of the meaning of a whole compound. Consider the verb *nokoru* 'remain, stay' (vi.), which expresses its literal meaning of 'remain' or 'exist' when used as the V2 of such compound verbs as *iki-nokoru* [live-remain], paraphrasable as *iki-te nokoru* '(a person) remain alive, survive', and *saki-nokoru* [bloom-remain], paraphrasable as *sai-te* (or *saita mama*) *nokoru* '(flowers) remain in bloom'. Because both V1 and V2 maintain their literal meanings, these examples are classified as thematic compound verbs. The same verb *nokoru*, however, has a radically different usage in such examples as *ure-nokoru* [be.sold-remain] 'remain unsold' and *yake-nokoru* [burn-remain] 'remain unburned'. These compounds are not subject to the simple paraphrase formula 'V1-*te* V2' but instead require the negative forms of V1 in paraphrase, as in *ure-zu ni nokoru* [be.sold-NEG DAT remain] for *ure-nokoru* 'remain unsold' and *yake-zu ni nokoru* [burn-NEG DAT remain] for *yake-nokoru* 'remain unburned'. The negative meaning in these examples is attributed to the aspectual meaning of *nokoru* that designates

incompletion of the event in V1, namely, 'The event in V1 remains (to be finished)'. This latter usage of *nokoru* is therefore classified as an L-aspectual verb. Apparently, the semantic shift from the physical meaning of 'remain, exist' to the aspectual meaning of 'incomplete unfolding' is not complete. *Syoohin ga ure-nokot-ta* [goods NOM sell-remain-PST] 'Goods remained unsold' implies not only that the selling event finished incompletely but also that the unsold goods remain at hand. This suggests that the L-aspectual *nokoru* somehow retains its physical meaning of 'remain'. This kind of double meaning is reminiscent of the reduced complement type of syntactic compound verbs (Section 3.2). As a matter of fact, the L-aspectual *nokoru*, which is intransitive, has its transitive counterpart *nokosu* 'leave' in the syntactic reduced-complement class (see Table 4), as in *(ryoori o) tabe-nokosu* 'leave (some food) uneaten'. *Nokoru* and *nokosu* are thus an intriguing pair whose members are distributed over the lexical and syntactic domains of aspectual compound verbs.

4.3 Lexical thematic compound verbs

We now turn to the class of lexical thematic compound verbs, in which the two members in V1 and V2, equipped with their own argument structure and literal meaning, jointly determine the argument relation and meaning of an entire compound verb.

Since the late 1970s, the internal semantic relations of compound verbs have been scrutinized by various researchers with diverse orientations (Nagashima 1976; Tagashira 1978; Morita 1990; Michiaki Saito 1992; Kageyama 1993; Yumoto 2005; Matsumoto 1996, 1998; Himeno 1999; Toratani 2002; Fukushima 2005). While these works tend to pursue detailed classifications of the internal semantic relations, we noted earlier that the semantic interpretation of thematic compound verbs is provided by the simple formula of 'V1-*te* V2' [V1 and V2]. Given this, an analysis of the internal semantic relations of thematic compounds is tantamount to the delimitation of the possible range of the semantic or pragmatic relationships that can hold between the two events in V1 and V2. The attempt at the delimitation will be aided by two factors. One is the iconic relation with the temporal sequence in which V1's event takes place prior to or simultaneously with V2's event. Another factor is what we might call "semantic coherence". Since a compound verb is a word, the whole compound must name a single coherent notion that is composed of two sub-events affecting each other in an organic manner. In *Watasi wa hataraki-tukare-ta* [I TOP work-get.tired-PST] 'I got tired from working too hard', for example, my fatigue must result directly from my hard work. These conditions naturally lead to a well-defined set of possible semantic relations between V1 and V2, as shown in (36) (Kageyama 1993; Matsumoto 1998; Yumoto 2005).

(36) a. Means of action: V2 BY V1-*ing*   [COMMON]
    uti-otosu [shoot-make.fall] 'shoot down', *yobi-atumeru* [call-gather]
    'call together', *nui-awaseru* [sew-put.together] 'sew together'

   b. Manner of motion or change: V2 WHILE V1-*ing*   [COMMON]
    nagare-otiru [flow-fall] 'fall, flowing', *hai-agaru* [crawl-go.up]
    'go up crawling', *tobi-noru* [jump-go.onto] 'jump onto'

   c. Cause-result: V2 FROM V1-*ing*   [LIMITED]
    syaberi-tukareru [speak-get.tired] 'get tired from talking too much',
    yoi-tubureru [get.drunk-fall.apart] 'get dead drunk'

   d. Purpose: V2 WITH A VIEW TO V1-*ing*   [RARE]
    kaburi-tuku [bite-attach] 'bite at', *daki-tuku* [hug-attach] 'throw oneself
    into another's arms'

   e. Coordination of synonymous verbs (dvandva): V1 and V2   [LIMITED]
    koi-kogareru [adore-pine.for] 'pine for', *odoroki-akireru* [be.surprised-
    be.stunned] 'be completely stunned'

Of these, Cause-result (36c) is restricted to a handful of designated head verbs. Purpose (36d) might be subsumed under Manner of motion (36b), and Coordination or dvandva (36e) tends to be reinterpreted as a kind of modification where V1 reinforces the meaning of V2. As a consequence, Means (36a) and Manner (36b) constitute the two major patterns, the choice of which is largely inferable from the event types of V2s. If V2 denotes a causative change of state, V1 will modify it most probably in a means relation (cf. Ishii 1983), whereas if V2 is a change-of-state or motion verb, V1 is likely to represent the manner of the motion or change.

　　In the previous literature, opinions diverge as to how many separate relations should be postulated and how they are formally represented. Researchers in the descriptive approach (Himeno 1999, to single out only one) tend to give overly fine-grained classifications based on actual examples attested in a variety of texts, whereas researchers in theoretical frameworks attempt to delineate only a limited number of core relations while relegating fine-grained differences to pragmatics in context (for example, Kageyama 1993 and Yumoto 2005 in Conceptual Semantics; Matsumoto 1996 in Lexical Functional Grammar; Toratani 2002 in Role and Reference Grammar; Fukushima 2005 in Dowty's macro-role semantics). Since the internal semantic relations of thematic compound verbs are readily read off from the meanings and event types of the two members, they need not be specified as rigid regulations. A plausible suggestion would be that the core relationships are captured in a rather simple manner by supplying V1's semantic structure into an open slot of one of the manner components of V2's semantic structure. The schematic representations in (37a, b) will give a rough idea of how the semantic structure of V1 is integrated into that of V2 to give rise to a coherent semantic structure of a whole compound verb.

(37) a. Manner relation
  korogari-otiru [roll-fall] 'fall rolling', i.e. 'roll down'
  V2 *otiru* 'fall': [x MOVE<Manner α> DOWN]
           ⇑ supplementation
     V1 *korogari* 'roll': *<Rotating>*

  b. Means relation
  hiki-ageru [pull-raise] lit. 'raise by pulling', i.e. 'pull up'
  V2 *ageru* 'raise': [x ACT<Manner α> ON y] CAUSE [y MOVE UP]
         ⇑ supplementation
    V1 *hiki* 'pull': *<Pulling>*

The manner and means relations are productive because insertion of V1's meaning into an open slot in V2's semantic structure is optimally simple. On the other hand, the cause and purpose relations (36c, d) are not so common because these meanings could be obtained only by invoking semantic information that is traditionally considered "pragmatic" – the kind of information that is contained in the Agentive quale (for the cause relation) or the Telic quale (for the purpose relation) of V2's qualia structure in the sense of Pustejovsky's (1995) Generative Lexicon theory.

No less important than the integration of the two events in V1 and V2 is the issue of how the argument properties (i.e. argument structures, case relations, and selection restrictions) of the two members are unified and realized in syntactic structure. Without delving into details of formal mechanisms (e.g. Matsumoto 1996; Gamerschlag 2001; Toratani 2002; Yumoto 2005; Fukushima 2005), only the representative patterns will be reviewed here, following Kageyama (1993).

Three major types of argument unification can be discerned. The first type is the simplest one where the two component verbs have exactly the same argument properties. The sentence *Watasi wa doa o keri-ake-ta*. 'I kicked the door open', for example, can be faithfully paraphrased as *Watasi wa doa o ket-te (doa o) ake-ta* 'I kicked the door and opened it', where V1 'kick' and V2 'open (vt.)' share the same subject and the same object. In this type of argument unification, which holds for the majority of transitive compound verbs, the argument properties are first identified between the two component verbs, and those of the head verb (V2) are projected to an entire compound verb, as schematically shown in (38).

(38) Pattern 1: The argument properties of two verbs are completely identical.
  [Example] *Watasi wa doa o keri-ake-ta*. 'I kicked the door open'

    keri-akeru 'kick open' $<Agent_{1=2}, Theme_{1=2}>$

 keri- $<Agent_1, Theme_1>$ *akeru* $<Agent_2, Theme_2>$
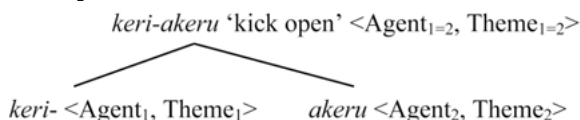
The second, less common pattern of argument unification is found with compound verbs where the selectional restrictions of the two component verbs differ although both share the same argument structure. In *sime-korosu* [choke-kill] 'choke to death', for example, V1 *sime-* 'choke' by itself selects nouns like 'throat' and 'neck' as its object (*niwatori no kubi o simeru* lit. 'choke the chicken's throat'), whereas V2 *korosu* by itself selects animate nouns as its object (*niwatori o korosu* 'kill the chicken', and not **niwatori no kubi o korosu* lit. 'kill the chicken's throat'). When the two verbs are combined, the whole compound selects an animate noun rather than 'throat' as its object: *niwatori (*no kubi) o sime-korosu* [chicken (*GEN throat) ACC choke-kill] 'choke the chicken to death'. This situation obtains only if the two object nouns are in a whole-part relation, with V1's object denoting a part of V2's object noun.

(39) Pattern 2: The argument properties of two verbs are partially identical.
     [Example] *Kare wa niwatori o sime-korosi-ta.* 'He choked the chicken to death.'

<div style="text-align:center">

sime-korosu 'choke to death' $\langle \text{Agent}_{1=2}, \text{Theme}_2 \rangle$

sime- $\langle \text{Agent}_1, \text{Theme}_1 \rangle$ *korosu* $\langle \text{Agent}_2, \text{Theme}_2 \rangle$

: where Theme$_1$ denotes a part of Theme$_2$.

</div>

An analogous instance is *arai-nagasu* [wash-let.flow] 'wash away', as in *kami-no yogore o arai-nagasu* 'wash away the grime from the hair', where the noun 'hair' corresponds to the object of V1 *arai-* 'wash' and the noun *yogore* 'grime' to that of V2 *nagasu* 'let flow, remove', thereby yielding the interpretation 'You wash your hair and remove the grime on it'.

The third pattern of argument unification, which is rather rare, takes place when the two component verbs select different semantic roles in the objects (or adjuncts). For example, the verb *kataru* 'talk' selects the topic of a talk as its object (*zinsei o kataru* 'talk about one's life'), while the verb *akasu* 'pass (the night)' selects an accusative-marked path phrase designating a night, as in *yoru o akasu* 'stay up all night'. Since the topic and the night bear different semantic roles, both of them may be realized syntactically, although the simultaneous occurrence of two accusative-marked phrases renders the sentence less than perfect, due to the so-called double-accusative constraint.

(40) Pattern 3: Two verbs have different internal arguments.
     [Example] *Hutari   wa natu   no yoru o, zinsei o   katari-akasi-ta.*
               two.people TOP summer GEN night ACC life   ACC talk-pass-PST
               'The two talked a summer night away about their life.'

<div style="text-align:center">

katari-akasu $\langle \text{Agent}_{1=2}, \text{Time}, \text{Theme} \rangle$

katari $\langle \text{Agent}_1, \text{Theme} \rangle$ *akasu* $\langle \text{Agent}_2, \text{Time} \rangle$

</div>

Interestingly, the two accusative-marked phrases need be aligned as shown in the example of (40), where V1's object *zinsei* 'one's life' is adjacent to the compound verb. Reversal of the two accusative phrases results in an ungrammatical sentence in (41), perhaps due to a perceptual disruption arising from the crossing of association links (Kageyama 1993).

(41) *zinsei*   *o*   *natu*   *no*   *yoru*   *o*   *katari-akasu*
      life   ACC   summer   GEN   night   ACC   talk-pass

Pattern 3 also applies to cases where V2 has an argument that is lacking in V1. For example, in *Ikada ga sima ni nagare-tuku* [drift-arrive] 'A raft drifts to an island', V1 *nagare-* 'drift' designates only a manner of motion ('being afloat') and the goal phrase ('to an island') is supplied by V2 *tuku* 'arrive'. A peculiar case is found with the bound root *-komu* 'move into, put into', which can be compounded with any kind of verb in V1, regardless of its transitivity and agentivity. In *poketto ni kane o osi-komu* 'shove the money into the pocket', for example, V1 *osi-* 'shove' by itself takes a Theme object ('money'), and V2 *-komu* the Goal phrase 'into the pocket'.

While the three patterns presented above apply to the vast majority of lexical thematic compound verbs, idiosyncratic exceptions remain which do not fall in any of them. A case in point is *yuzuri-ukeru* [bestow-receive] 'have something handed over from someone', where the 'receiver' of V2 *ukeru* 'receive' is realized as the subject and the 'giver' of V1 *yuzuru* 'bestow' is embodied as an ablative phrase with *kara* 'from'. See Fukushima (2005) for discussion on this phenomenon.

4.4 Combinatory restrictions on lexical compound verbs

At a pre-theoretical level, Japanese grammarians sometimes noticed the strong tendency for a transitive verb to be compounded with another transitive verb, and for an intransitive verb to be compounded with another intransitive verb. The "transitive parity" as Jacobsen (1992) calls it thus predicts that the transitive verb *nagasu* 'let flow' is compounded with the transitive verb *otosu* 'drop' to make up the transitive compound *nagasi-otosu* [let.flow-make.fall] 'wash (dirt) off', whereas the intransitive *nagareru* 'flow' is combined with the intransitive *otiru* 'fall' to form the intransitive compound *nagare-otiru* 'flow off'. Ungrammatical compounds result if a transitive V1 is combined with an intransitive V2 (**nagasi-otiru* '(dirt) washes off') or an intransitive V1 is combined with a transitive V2 (**nagare-otosu* lit. 'remove by flowing'). Japanese grammarians have also observed a notable tendency that V1s are

likely to represent an agent's action, and V2s change of state or change of location (Ishii 1983; Hayatsu 1989).

As attractive as it is, transitive parity is too weak to capture the nature of the combinatory restrictions – too weak partly because it cannot account for well-formed combinations of transitive and intransitive verbs, such as *(me o) naki-harasu* [cry (vi.)-make.swell (vt.)] 'cry one's eyes out', where V1 *naku* 'cry' is intransitive and V2 *harasu* 'make swell' is transitive, and *(kagi o) sagasi-mawaru* [look.for (vt.)-go.round (vi.)] 'look around for (the key)', where V1 *sagasu* 'look for' is transitive and V2 *mawaru* is intransitive, and partly because it cannot rule out supposedly well-formed combinations of two intransitive verbs that are actually judged unacceptable, as in *(me ga) *naki-hareru* [cry (vi.)-swell (vi.)] 'One's eyes are swollen because of crying'. An interesting pair is *suberi-otiru* 'slip off', which is a non-volitional intransitive verb, and *suberi-oriru* 'ski down (a slope)', which is a volitional intransitive verb. While V1 *suberi-* 'slide, slip' in itself is ambiguous between an accidental slip (an unaccusative verb) and a volitional sliding (an unergative verb), V2s *otiru* 'fall' and *oriru* 'go down, climb down' are clearly differentiated by the subject's volitionality in the downward motion, *otiru* being non-volitional (i.e. an unaccusative verb) and *oriru* being volitional (i.e. an unergative verb). A more accurate generalization should thus be sought in the volitional/accidental dimension of meaning. Typically, transitive and unergative verbs take a volitional agent as subject, whereas unaccusative verbs select a non-volitional subject. These considerations led Kageyama (1993, 1999) to postulate a generalization dubbed the "Transitivity Harmony Principle".

(42) The Transitivity Harmony Principle
    Given the three argument structure schemas below, lexical compound verbs are built by combining two verbs of the same type of argument structure.
    (a) transitive verbs (Tr.): (Agent <Theme>)
    (b) unergative verbs (Unerg.): (Agent < >)
    (c) unaccusative verbs (Unacc.): <Theme>

The sameness of argument structure type stipulated in (42) is determined by the presence or absence of an external argument (typically, agent or experiencer). Transitive and unergative verbs, equipped with an external argument, count as the same type whereas unaccusative verbs, lacking an external argument, form a disparate type. The Transitivity Harmony Principle thus allows the combinations of "Tr + Tr", "Tr + Unerg", "Unerg + Tr", "Unerg + Unerg", and "Unacc + Unacc", while at the same time ruling out the combinations of "Tr + Unacc", "Unacc + Tr", "Unerg + Unacc", and "Unacc + Unerg". As predicted, the examples in (43) are judged well-formed (and in fact, commonly used) whereas such examples as those in (44) are disallowed.

(43) a. transitive V1 + unergative V2
          (teki o) mati-kamaeru [(enemy ACC) wait-be.prepared] 'brace oneself for
          the enemy'

   b. unergative V1 + transitive V2: *warai-tobasu* [laugh-fly] 'laugh away'

   c. unaccusative V1 + unaccusative V2: *sitatari-otiru* [drip-fall] 'drip down'

(44) a. *transitive V1 + unaccusative V2
          **tuki-otiru* [push-fall] 'push and fall'
          Cf. tr.+ tr.: *tuki-otosu* [push-make.fall]

   b. *unaccusative V1 + transitive V2
          **ore-mageru* [be.bent-be.twisted] 'be bent and twisted'
          Cf. tr.+tr.: *ori-mageru* [bent-twist]

   c. *unergative V1 + unaccusative V2
          *(me ga) *naki-hareru* [(eyes NOM) cry-get.swollen] '(eyes) swell because of
          crying hard'

   d. *unaccusative V1 + unergative V2
          **koroge-oriru* [tumble-step.down] 'tumble and step down'
          Cf. unacc.+unacc.: *koroge-otiru* [tumble-fall]

The Transitivity Harmony Principle has since been the subject of a number of arguments pro and con (Kageyama 1996; Matsumoto 1998; Nishiyama 1998; Yumoto 2005; Zhu 2009; Chen 2010; Mamoru Saito 2013; cf. also Naumann and Gamerschlag 2003), and alternative generalizations have been proposed such as the "same subject condition" (Matsumoto 1998), which requires the subjects of two component verbs to be identical, the "priority of unaccusativity" (Yumoto 2005), and the "matching principle" of macro-roles (Fukushima 2005). Nishiyama (2008) reviews the technical details of these competing analyses and concludes that more research is needed to decide which condition captures the facts in an optimal way. Here I will continue to elaborate on the merits of the Transitivity Harmony Principle.

First, the empirical validity of this principle is supported by statistical data. In my database of about 2,750 lexical compound verbs (Kageyama and Kanzaki 2014), approximately 84% obey the Transitivity Harmony Principle. Moreover, a significant disparity is observed between aspectual and thematic compound verbs: Almost 90% of the thematic compound verbs comply with the principle, while only 70% of the aspectual compound verbs do so. This discrepancy is a natural consequence of the functional difference between the two classes of compounds. The thematic compound verbs must obey the Transitivity Harmony Principle faithfully because their event integration depends on the identity of the argument structure types between V1 and V2. The aspectual compound verbs do not necessarily follow this principle because their argument structure is basically determined by V1s alone.

Second, the Transitivity Harmony Principle has a functional motivation. Take the compound verb *humi-tubusu* 'trample down', for example. In order for this compound to make up a single coherent event of someone's crushing something by trampling it, the agent of V1 must necessarily be identical to the agent of V2. Were the V1 represented by a transitive verb and the V2 by an unaccusative verb, then the V1's agent could not exert its control over the V2's spontaneous event, thus yielding an ungrammatical compound like *humi-tubureru [trample-be.crushed]. Consequently, the Transitivity Harmony Principle can be construed as a regulation on "agent sharing" in event unification: V1 and V2 must share one and the same agent if they have an agent at all.

As mentioned above, approximately 90% of the thematic compound verbs comply with the Transitivity Harmony Principle. What about the exceptional 10%? Almost all of the exceptional cases have the supposedly ungrammatical combinations of "Tr + Unacc" and yet are used more or less commonly. Kageyama (forthcoming) develops a principled analysis that accounts for such deviant cases as resulting from semantic reanalysis of the transitive-transitive combinations. In *Zisin de zimen ga moti-agat-ta.* [earthquake INS ground NOM hold(vt.)-go.up(vi.)] 'The ground lifted up because of an earthquake', for example, the meaning of V1 *motu*, literally meaning 'hold in one's hand', is semantically reanalyzed as designating a natural force (earthquake) that is responsible for the ground's upheaval. On this analysis, since the human agent is obliterated from V1's argument structure, only the identification of theme arguments suffices and hence the superficial combination of transitive and unaccusative verbs is permitted. As predicted, the Transitivity Harmony Principle is more easily violated with lexical aspectual compound verbs, as exemplified by *kaki-ageru* [write (tr.)-finish (tr.)] 'finish writing' and *kaki-agaru* [write (tr.)-finish (intr.)] 'be done, writing' (cf. Kageyama, forthcoming).

5 Conclusion and future research perspectives

This chapter has surveyed the characteristic properties of V-V compound verbs. The major division is drawn between lexical and syntactic compound verbs, with the term "lexical" referring to the fact that two lexical verbs are directly compounded, and the term "syntactic" to the fact that the head verb selects a syntactic clause as its complement. Syntactic compound verbs are divided into three subclasses, depending on the type of the complement clause the head verb selects, whereas lexical compound verbs are partitioned into two subclasses, thematic and aspectual. The subclasses and their mutual relationships are summarized in Table 6.

Table 6: Subclasses and mutual relationships of V-V compound verbs

| Class | Lexical | | syntactic | | |
	Thematic	Aspectual	V' complement	VP with control	VP with raising
I. Structure	left-to-right modification	complementation/ right-to-left modification	comple-mentation	comple-mentation	comple-mentation
II. Meaning	thematic	aspectual (or thematic-aspectual)	thematic-aspectual	aspectual	aspectual
III. Transitivity Harmony	Applies with almost absolute accuracy.	Applies with reasonable accuracy.	Irrelevant.	Irrelevant.	Irrelevant.
IV. Desemanti-cization	V1	V2	V2	V2	V2
V. Hierarchical relation	lowest ←——————————————————————————→ highest				

Row IV, desemanticization, shows which of V1 and V2 is likely to undergo semantic bleaching. In lexical thematic compound verbs, the first verb (V1) is amenable to loss of its original meaning caused by phonetic erosion, thus turning to a prefix, as in *hiQ-* in *hip-pagasu* 'take off violently' from *hiki-hagasu* [pull-take. off] and *buQ-* in *buk-kowasu* 'demolish' from *buti-kowasu* [knock-break] (cf. Vance 2001). In all the other types, V2s have undergone desemanticization while V1s maintain their original meanings and argument structure.

Table 6 thus shows a bird's-eye view of the system of Japanese V-V compound verbs, pinpointing not only how different the five classes of V-V compound verbs are from each other, but also how similar or how "continuous" they are to each other. Evidently, the class of lexical aspectual compound verbs plays a vital role in bridging the lexical and syntactic domains in the whole inventory of Japanese V-V compound verbs. These intricate similarities and differences ranging from morphological to syntactic and semantic properties could not be given a principled explanation either in extreme lexicalist frameworks or in extreme syntactic frameworks such as Distributed Morphology (Nishiyama 1998, 2008) or Minimalist Syntax (Mamoru Saito 2013). Developing the idea of Shibatani and Kageyama (1988), Kageyama (1993) proposed a model of Modular Morphology in which word formation is allowed to take place in various modules of grammar (lexicon, syntax, and phonology), and Yumoto (2005) further pursued the validity of this model. Fruitful outcomes are expected to emerge only from those studies that are directed toward elucidating the nature of the similarities and differences among different types of compounding with balanced attention paid to all of morphology, syntax, and semantics.

Acknowledgments

References

Baker, Mark. 1989. Object sharing and projection in serial verb constructions. *Linguistic Inquiry* 20. 513–553.

Bloch, Bernard. 1946. Studies in Colloquial Japanese I: Inflection. *Journal of the American Oriental Society* 66. 97–109. Reprinted in Roy A. Miller (ed.), 1969. *Bernard Bloch on Japanese*. New Haven: Yale University Press.

Chen, Jie-Yi. 2010. Goiteki fukugōdōshi no jitakōtai to gokeisei [Transitivity alternation and word formation of lexical compound verbs]. *Nihongo Bunpō* [Journal of Japanese Grammar] 10(1). 37–53. Tokyo: Kurosio Publishers.

Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.

Frellesvig, Bjarke, Stephen Horn, Kerri Russell and Peter Sells. 2010. Verb semantics and argument realization in pre-modern Japanese: A preliminary study of compound verbs in Old Japanese. *Gengo Kenkyu* 138. 25–65.

Fukui, Naoki, Shigeru Miyagawa and Carol Tenny. 1985. Verb classes in English and Japanese: A case study in the interaction of syntax, morphology and semantics. *Lexicon Project Working Papers* 3. Center for Cognitive Science, MIT.

Fukuda, Shin. 2012. Aspectual verbs as functional heads: Evidence from Japanese aspectual verbs. *Natural Language & Linguistic Theory* 30(4). 965–1026.

Fukushima, Kazuhiko. 2005. Lexical V-V compounds in Japanese: Lexicon vs. syntax. *Language* 81. 568–612.

Gamerschlag, Thomas. 2001. Complex predicate formation and argument structure of Japanese V-V compounds. In Noriko M. Akatsuka and Susan Strauss (eds.), *Japanese/Korean Linguistics* 10. 532–544. Stanford: CSLI Publications.

Harada, S. I. 1976. Honorifics. In Masayoshi Shibatani (ed.), *Syntax and semantics* 5: *Japanese generative grammar*, 499–561. New York: Academic Press.

Hayatsu, Emiko. 1989. Yūtsui-tadōshi to mutsui-tadōshi no chigai nitsuite [On the differences between paired and unpaired transitive verbs]. *Gengo Kenkyu* 95. 231–256.

Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Himeno, Masako. 1999. *Fukugōdōshi no kōzō to imi yōhō* [The structure and meaning/usage of compound verbs]. Tokyo: Hituzi Syobo.

Inoue, Kazuko. 1976. *Henkei bunpō to nihongo* [Transformational grammar and Japanese]. Tokyo: Taishukan Shoten.

Ishii, Masahiko. 1983. Gendaigo fukugōdōshi niokeru ichi-kanten [A look at compound verbs in contemporary Japanese]. *Nihongogaku* 2(8). 79–90.

Jacobsen, Wesley M. 1992. *The transitive structure of events in Japanese*. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1989. The place of morphology in the grammar: Verb-Verb compounds in Japanese. In Geert Booij and Jaap van Marle (eds.), *Yearbook of morphology* 2, 73–94. Dordrecht: Foris.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1999. Word formation. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 297–325. Oxford: Blackwell.

Kageyama, Taro. 2013. Goiteki fukugōdōshi no shin taikei [A new system of lexical compound verbs]. In Taro Kageyama (ed.), *Fukugōdōshi kenkyū no saizensen* [New explorations into the mysteries of compound verbs], 3–46. Tokyo: Hituzi Syobo.

Kageyama, Taro. this volume. Chapter 14: Lexical Integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. forthcoming. Agents in anticausative and de-causative compound verbs. In Taro Kageyama and Wesley M. Jacobsen (eds.), *Valency and transitivity alternations: Studies on Japanese and beyond*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Kyoko Kanzaki. 2014. Compound verb lexicon (online database). Tokyo: National Institute for Japanese Language and Linguistics. http://vvlexicon.ninjal.ac.jp/

Kishimoto, Hideki. 2006. Japanese syntactic nominalization and VP-internal syntax. *Lingua* 116. 771–810.

Kishimoto, Hideki. 2013. Tōgoteki fukugōdōshi no kaku to tōgososei [Case and syntactic features of syntactic compound verbs]. In Taro Kageyama (ed.), *Fukugōdōshi kenkyū no saizensen* [New explorations into the mysteries of compound verbs], 143–183. Tokyo: Hituzi Syobo.

Koizumi, Masatoshi. 1999. *Phrase structure in Minimalist syntax*. Tokyo: Hituzi Syobo.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu. 1983. *Shin nihon bunpō kenkyū* [New studies on Japanese grammar]. Tokyo: Taishukan Shoten.

Kuroda, S.-Y. 1979. *Generative grammatical studies of the Japanese language*. New York: Garland.

Martin, Samuel. E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Masica, Colin P. 1976. *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.

Masuoka, Takashi. 1987. *Meidai no bunpō* [The grammar of propositions]. Tokyo: Kurosio Publishers

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI and Tokyo: Kurosio Publishers.

Matsumoto, Yo. 1998. Nihongo no goiteki-fukugōdōshi niokeru dōshi no kumiawase [The combinatory possibilities in Japanese V-V lexical compounds]. *Gengo Kenkyu* 114. 37–83.

Miyagawa, Shigeru. 1989. Syntax and semantics 22: *Structure and case marking in Japanese*. New York: Academic Press.

Morita, Yoshiyuki. 1990. *Nihongo gaku to nihongo kyōiku* [Japanese-language studies and Japanese-language teaching]. Tokyo: Bonjinsha.

Nagashima, Yoshio. 1976. Fukugōdōshi no kōzō [Structure of compound verbs]. In Takao Suzuki (ed.), *Nihongo no goi to hyōgen* [Japanese vocabulary and expressions], 63–104. Tokyo: Taishukan Shoten.

Nakatani, Kentaro. this volume. Chapter 11: Complex predicates with -*te* gerundive verbs. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Naumann, Ralf and Thomas Gamerschlag. 2003. Constraining the combinatorial patterns of Japanese V-V compounds: An analysis in dynamic event semantics. *Journal of Semantics* 20(3). 275–296

Nishigauchi, Taisuke. 1993. Long distance passive. In Nobuko Hasegawa (ed.), *Japanese syntax in comparative grammar*, 79–114. Tokyo: Kurosio Publishers.

Nishiyama, Kunio. 1998. V-V compounds as serialization. *Journal of East Asian Linguistics* 7. 175–217.

Nishiyama, Kunio. 2008. V-V compounds. In Shigeru Miyagawa and Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 320–347. Oxford: Oxford University Press.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Saito, Mamoru. 2013. Conditions on Japanese phrase structure: From morphology to pragmatics. *Nanzan Linguistics* 9. 119–145. Nagoya: Nanzan University.

Saito, Michiaki. 1992. *Gendai nihongo no gokōseiron teki kenkyū: Go ni okeru katachi to imi* [A Study of the word composition in contemporary Japanese: Form and meaning of words]. Tokyo: Hituzi Syobo.

Shibatani, Masayoshi. 1973. Where morphology and syntax clash: A case in Japanese aspectual verbs. *Gengo Kenkyu* 64. 65–96.

Shibatani, Masayoshi. 1978. *Nihongo no bunseki: Seisei bunpō no hōhō* [An analysis of Japanese: The method of generative grammar]. Tokyo: Taishukan Shoten.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar: Postsyntactic compounds in Japanese. *Language* 64. 451–484.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Tagashira, Yoshiko. 1978. *Characterization of Japanese compound verbs*. Chicago, IL: University of Chicago dissertation.

Tamaoka, Katsuo, Hyunjung Lim and Hiromu Sakai. 2004. Entropy and redundancy of Japanese lexical and syntactic compound verbs. *Journal of Quantitative Linguistics* 11(3). 233–250.

Toratani, Kiyoko. 2002. *The morphosyntactic structure and logical structures of compound verbs in Japanese*. Buffalo, NY: State University of New York dissertation.

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. Albany: State University of New York Press.

Vance, Timothy J. 2001. Semantic bifurcation in Japanese compound verbs. In Noriko M. Akatsuka and Susan Strauss (eds.), *Japanese/Korean Linguistics* 10. 365–377. Stanford: CSLI Publications.

Yumoto, Yoko. 2001. A modular approach to the semantic interpretation of syntactic V-V compounds. *Journal of Japanese Linguistics* 17. 125–150.

Yumoto, Yoko. 2005. *Fukugōdōshi haseidōshi no imi to tōgo* [Semantics and syntax of compound verbs and complex verbs]. Tokyo: Hituzi Syobo.

Yumoto, Yoko. 2009. Modularity of word formation: Differences between two types of Japanese compound verbs. In Hiroto Hoshi (ed.), *The dynamics of the language faculty*, 203–230. Tokyo: Kurosio Publishers.

Zhu, Chunri. 2009. Fukugōdōshi no jita-taiō nitsuite [Concerning the transitive-intransitive correspondences in compound verbs]. *Sekai no nihongo kyōiku* [Japanese-language education in the world] 19. 89–106. Tokyo: Bonjinsha.

Yoko Yumoto
9 Conversion and deverbal compound nouns

1 Introduction

This chapter deals with issues involving verb-to-noun (V-to-N) conversion and compounding based on it. In Japanese, V-to-N conversion takes place typically with the infinitive (*ren'yō*) forms of native verbs, which function as nouns without any special suffix, as in *sasae* 'support' from the infinitive *sasae-* 'to support' (cf. the present form *sasae-ru*) and *turi* 'fishing, angling' from the infinitive *turi-* 'to fish, angle' (cf. the present form *tur-u*). Deverbal nouns created by conversion are then subject to compounding with another noun (or adjective) on their left that serves as their argument, as in *hige-sori* [beard-shave] 'shaver, razor' (*-sori* < *sor-u* 'to shave'), or as their adjunct, as in *sumibi-yaki* [charcoal-broil] 'charbroil' (*-yaki* < *yak-u* 'to broil'). Since Japanese deverbal compounds of the form "N + deverbal N" are partially similar to but even more complex than the English "deverbal" or "synthetic" compounds marked with the suffixes such as *-ing* (e.g. *wind-surfing*) and *-er* (e.g. *taxi-driver*) or with no overt suffix (e.g. *Google search*), they present a number of intriguing properties as regards their meanings and syntactic usages.

The deverbal compounds treated in this chapter superficially resemble N-V compound verbs such as *kosi-kakeru* [hips-hang] 'to seat oneself'. Specifically, they share similarities in the internal relations of the two constituents. Just as the N-V compound verb *kosi-kakeru* 'to seat oneself' is analyzed as having an object-verb relation between the noun *kosi* 'hips' and the verb *kakeru* 'hang, put', so is the deverbal compound noun *kane-moti* 'rich man' analyzed as having the same relation between *kane* 'money' on the left and *moti* (<*mot-u* 'hold) on the right. As to their syntactic status, however, the two types of compounds are radically distinct: N-V compound verbs are capable of taking tense inflection directly because they are verbs, while the deverbal compounds are not because they are nouns (see Chapter 7 [Kageyama, this volume]). The nominal character of deverbal compounds is also reflected in the diversity of their meanings, which range from concrete and abstract entities to actions, events, and states.

This chapter is organized as follows. Section 2 introduces the basic properties of V-to-N and A-to-N (adjective-to-noun) conversion, while the subsequent sections are devoted to the major types of deverbal compounds, discussing the mechanisms of semantic interpretation. Special attention will be paid to three problems: (i) how the internal relations of deverbal compound nouns can be captured (Section 3), (ii) how deverbal compounds as a whole take on a variety of meanings without the aid of any morphological clues (Section 4), and (ii) how some deverbal compound nouns are further subject to complex predicate formation with *suru* 'do' (Section 5). Section 6 gives a summary and points to a future direction of research.

2 General properties of V-to-N conversion

This section outlines the morphophonological and semantic characteristics of conversion, a word formation process that changes the lexical category of a word without a particular affix. Unlike languages like English and French, where conversion processes are bidirectional from N to V (e.g. *a table*$_N$ → *to table*$_V$) and V to N (e.g. *catch*$_V$ → *a catch*$_N$), conversion in Japanese is unidirectional, being limited to change from V to N and some other categories. N-to-V conversion is not allowed because the categorial change of N to V calls for a morphological verbalizer like *-r* or the light verb *suru* 'do' (cf. Chapter 1 [Kageyama and Saito, this volume]).

In Japanese, two inflecting categories of the native stratum – verbs (V) and adjectives (A) – can be converted to non-inflecting categories, which include not only nouns but (to a limited extent) conjunctions, particles and adverbs as well. Table 1 summarizes possible types of conversion in Japanese.

Table 1: Types of conversion

Category change	Examples
V → N	*obi* 'sash, wide belt' (< *obi ru* 'wear'), *osi* 'boldness' (< *os u* 'push'), *osore* 'fear' (< *osore ru* 'fear'), *kangae* 'thought, idea' (< *kangae ru* 'think'), *toori* 'street' (< *toor u* 'go past'), *ari* 'existence, presence' (< *ar u* 'be, exist')
V → CONJCT	*oyobi* 'and' (< *oyob u* 'extend to'), *tinami ni* 'in this connection' (< *tinam u* 'be connected with')
Complex particles	*X ni tuki* 'per X' (< *ni tuk u* 'attach to X'), *X o hazime* 'including X' (< *o hazime ru* 'begin X'), *X ni oite* 'at' (<*ok u* 'put on X')
V → ADV	*amari* '(not) much' (< *amar u* 'be left over'), *bakari* 'about' (< *hakar u* 'measure')
A → N	*nasi* 'none, nothing, absence' (< *na i* 'null, non-existent'), *susi* 'sushi' (< *su i* 'sour'); *asi baya* 'quick steps' (*asi* 'leg', *haya i* 'fast')

As shown in Table 1, verbs undergo conversion in the *ren'yō* or "infinitive" form, whereas some adjectives undergo conversion in the conclusive form (*-si*) of classical Japanese and others in the form of a bare stem. For example, the coordinate compound *ari-nasi* 'presence or absence' is composed of *ari* 'presence', the infinitive form of the verb *aru* 'be, exist', and *nasi* 'absence', the classical conclusive form corresponding to the adjective *na-i* 'non-existent'. Nouns based on adjectives of the classical conclusive form are rare, while nouns based on the bare forms of adjectives are fairly common, especially when they are embedded in compound structure.

Concerning the function of *ren'yō* verbs, Ōno (1953) argues on historical and etymological grounds that the final vowel *-i* that appears after the stems of consonant-ending verbs, as in *sakeb-i* 'shout' (stem *sakab-* 'to shout') and *tur-i* 'fishing' (stem *tur-* 'to fish'), is a nominalizing suffix. Although its validity must be ascertained by historical data, this hypothesis would be hard to accept from a synchronic

point of view because the hypothesized *-i* suffix does not show up with vowel-ending stems like *mane* 'mimicry' (stem *mane-* 'to mimic') and *tasuke* 'help' (stem *tasuke-* 'to help') (cf. *mane-i, *tasuke-i). Moreover, the putative *-i* suffix has no place in the nominalization of adjectives, which are mostly nominalized in the form of a bare stem. In what follows, we chiefly discuss V-to-N conversion, referring to A-to-N conversion only when necessary.

2.1 Conversion of simplex verbs

We will start with noun conversion of simplex verbs. Relevant examples can be classified in terms of the transitivity of source verbs as shown in (1).

(1) a. Intransitive verbs (unaccusative and unergative):
        aki 'vacancy' (< *ak-u* 'to open'), *kaeri* 'the way back from somewhere'
        (< *kaer-u* 'to return'), *yure* 'shake' (< *yure-ru* 'to shake'), *warai* 'laughter'
        (< *wara(w)-u* 'to laugh'), *zure* 'gap' (< *zure-ru* 'to deviate')

     b. Transitive verbs:
        iyasi 'healing' (< *iyas-u* 'to heal'), *sasae* 'support, prop' (< *sasae-ru*
        'to support'), *seme* 'blame' (<*seme-ru* 'to blame'), *tikai* 'oath' (< *tika(w)-u*
        'to swear'), *utusi* 'copy' (< *utus-u* 'to copy')

     c. Ditransitive verbs:
        kasi 'rent, lending' (< *kas-u* 'to lend'), *kari* 'borrowing' (< *kari-ru*
        'to borrow'), *osie* 'teaching, lesson' (< *osie-ru* 'to teach'), *watasi* 'ferry'
        (< *watas-u* 'to transfer, ferry')

     d. Psychological verbs:
        akogare 'longing' (< *akogare-ru* 'to long for'), *akirame* 'abandonment'
        (< *akirame-ru* 'to abandon'), *tanosimi* 'pleasure' (< *tanosim-u* 'to enjoy'),
        yorokobi 'delight' (< *yorokob-u* 'be delighted')

Many of such deverbal nouns, especially when they are used as independent nouns, lack a semantically transparent correspondence with their source verbs, so much so that the etymological relationships may not be easily recognized by native speakers of today (Nishio 1988). Notable examples of semantically opaque deverbal nouns include *tatami* 'a *tatami* mat' (< *tatam-u* 'to put away'), *tati* 'sword' (< *tat-u* 'to cut off'), and *tayori* 'letter, news' (< *tayor-u* 'to rely'). Such deverbal nouns with specialized meanings must be registered one by one in the lexicon. Even those whose derivational relations with the source verbs are more or less transparent in modern Japanese are not fully productive. According to Takeo Miyake (referred to in Nishio 1988), only 24.1% of the commonly used basic verbs (about 482 out

of 2000) have the nominal usage of the *ren'yō* forms. The restricted productivity of V-to-N conversion gives rise to lexical gaps as exemplified by the nouns in (2) which are unacceptable as independent words.

(2)   a.   One-mora nouns: **mi* (< *mi-ru* 'to see'), **ki* (< *ki-ru* 'to wear', or *ku-ru* 'come'), **si* (< *su-ru* 'to do'), **ni* (< *ni-ru* 'to resemble',' to cook')

       b.   Two-mora nouns: **huki* (< *huk-u* 'blow'), **hai* (< *ha(w)-u* 'to crawl'), **tabe* (< *tabe-ru* 'to eat'), **tati* (< *tat-u* 'to stand')

       c.   Three-mora nouns: **wasure* (< *wasure-ru* 'to forget'), **nokosi* (< *nokos-u* 'to leave behind')

It should be stressed here that all of these forms become well-formed when put in compound structure, as will be discussed in Section 2.2.

In previous literature, such lexical gaps have been accounted for by recourse to two conditions, neither of which is entirely feasible, however. One is a phonological constraint on the number of moras that compose a deverbal noun. Nishio (1988) holds that deverbal nouns that consist of only one or two moras tend to be unacceptable because they lack phonological stability as independent words. Such a phonological condition, however, expresses only a tendency and meets with a number of counterexamples such as *de* 'gush' (< *de-ru* 'go out'; e.g. *(mizu no) de ga warui* [(water GEN) gush NOM bad] '(water) doesn't run well'), *ne* 'sleep' (< *ne-ru* 'sleep'; e.g. *ne ga tari-nai* [sleep NOM suffice-NEG] 'not sleep enough'), and *aki* 'vacancy' (< *ak-u* 'open'; e.g. *aki ga aru* [vacancy NOM exist] 'there is a vacancy').

The other condition conceivable is based on the principle of "lexical blocking" (also called "pre-emption") in morphology, which holds that creation of a new word by a rule is blocked by an existing word that has the same meaning. Following this condition, the unacceptability of the native deverbal nouns in (3a, b) below, for example, would be attributed to the existence of the synonymous Sino-Japanese nouns.

(3)   a.   **erabi* (< *erab-u* 'choose'), supposedly blocked by S-J noun *sentaku* 'choice'.

       b.   **usinai* (< *usina(w)-u* 'lose'), supposedly blocked by S-J noun *soositu* 'loss'.

It is known, however, that lexical blocking is not an absolute condition because it is based on the vague notion of "synonymy". In fact, it is not difficult to find pairs of a native deverbal noun and a Sino-Japanese noun that are more or less synonymous, such as *ne* (native) 'sleep' and *suimin* (S-J) 'sleep', or *aki* (native) 'vacancy' and *kuusitu* (S-J) 'vacant room' or *ketuin* (S-J) 'vacant post'.

Nishio (1988: 74–75) entertains another idea to explain the relatively low productivity of simplex deverbal nouns from a semantic standpoint. According to him,

the base verbs underlying them allow a wide range of meanings and hence are not specific enough when converted to nouns. Taken at its face value, Nishio's account has many exceptions. For example, the deverbal noun *watasi* 'ferry' (< *watas-u*) is well-formed despite the fact that the base verb *watasu* covers a wide range of meanings such as 'to hand, to give, to deliver, to relinquish' other than 'to ferry'. Likewise, the deverbal noun *de* (< *de-ru* 'go out') is fully acceptable with a specific meaning of 'one's turn to go on stage' even those the base verb *deru* can refer to any action of going out or appearing on the scene.

Kageyama (2001, 2002) suggests a novel account for the low productivity of V-to-N conversion and its general incapability of designating discrete entities in simplex words. His idea draws on a general observation that the Japanese language often lacks specificity or "boundedness", compared with English. In particular, Japanese verbs, even causative change-of-state verbs, tend to be non-bounded and non-committal to the entailment of resultant states, and this general tendency provides a motivation to add a specific morpheme as the head to a simplex deverbal noun, which will make it semantically bounded. Thus, *hoe 'a bark' (< *hoeru* 'to bark'), unacceptable in isolation, becomes acceptable if it is followed by a specific noun *koe* 'voice' as in *hoe-goe* [bark-voice] 'a bark'. Kageyama (2001) also draws attention to an important fact that has been overlooked in the previous literature on V-to-N conversion. The formation of new words is governed by one of two principles: (i) regular, rule-governed **productivity**; and (ii) sporadic, unsystematic **creativity** due to the speaker's free imagination (cf. Chapter 1 [Kageyama and Saito, this volume] and references cited therein). The former ("productivity") has to do with regular processes of word formation based on the general principles of morphology as well as on the inherent properties of particular affixation and compounding processes. The latter ("creativity"), on the other hand, is motivated by the speaker's need to create a new word in a particular context. That is, it is a process of naming an unknown entity by applying metaphorical, metonymic, or other figurative means, at random, to existing word formation processes. In this sense, conversion provides an ideal environment for morphological creativity because it involves no particular affix, leaving speakers with freedom to assign a new, creative meaning to a deverbal noun in a particular pragmatic context. For example, there would be no principled reason that a sauce to drip on food is called *tare* (< *tare-ru* 'to drip'). Though there are many things in the real world that "drip", the deverbal noun *tare* happens to be exploited to name culinary sauce. Such examples as *tatami* 'a *tatami* mat' (< *tatam-u* 'to put away') and *tati* 'sword' (< *tat-u* 'to cut off') which are semantically opaque also belong to creative coinage.

Thus the meanings of deverbal nouns due to morphological creativity are hard to predict. By contrast, we could and should expect to find an orderly system of semantic interpretation if we limit ourselves to the deverbal nouns formed by productivity, i.e. by rule-governed principles. Kageyama (2001) draws a fundamental distinction between "event nouns", which denote actions, events, states, etc., and

"entity nouns", which represent entities. Event nouns derived from verbs are exemplified in (4).

(4) Event nouns
- a. Action / Event: *oyogi* 'swim' (< *oyog-u* 'to swim'), *nemuri* 'sleep' (< *nemur-u* 'to sleep'), *odori* 'dance' (< *odor-u* 'to dance'), *arasoi* 'fight' (< *araso(w)-u* 'to fight'), *yure* 'shake, swing' (< *yure-ru* 'to shake')

- b. Manner or degree of action: *kire* 'sharpness (of a knife), dryness (of beer)' (< *kire-ru* 'to be cut off'), *nori* 'getting into the rhythm' (< *nor-u* 'to get on'), *suwari* 'fitness' (< *suwar-u* 'to sit')

- c. Resultant state: *bure* 'unstableness, blur' (< *bure-ru* 'not to hold firmly'), *yugami* 'warp' (< *yugam-u* 'to warp')

These nouns represent intangible notions like action, manner, and state. Among them, the most productive is the formation of action and event nouns in (4a). The manner or degree nouns in (4b) are often modified by evaluative adjectives such as *yoi* 'good' and *warui* 'bad', as in *Kono kyoku wa nori ga yoi* [this music TOP get.on NOM good] 'This music is easy to get into'. The examples in (4c) represent the resultant states implied by the base verbs.

The semantic properties of such deverbal nouns as those in (4) can be captured most plausibly by employing some kind of lexical semantic representation designed for analyzing the meaning of verbs. In the literature, there are numerous models of verb semantics ranging from conceptual semantics to cognitive semantics to formal semantics. For convenience' sake, we will use here a modified version of Kageyama's (1996) Lexical Conceptual Structure (LCS). In Kageyama's framework, LCS representations not only show the lexical meaning of a given verb but its aspectual type. The schemas in (5) illustrate how various aspectual types of eventualities are represented in LCS.

(5) a. Intransitive activity (e.g. *dance*): [x ACT IN MANNER α] (=x acts in a manner specified by α)

- b. Transitive activity (e.g. *wipe*): [x ACT ON y IN MANNER α] (=x acts on y in a manner specified by α (with no specific result))

- c. Achievement of change of state (e.g. *dry* intr.): [BECOME [y BE AT-[$_{STATE}$ α]]] (=y becomes α)

- d. Volitional motion (*run*-verbs): [x ACT] CAUSE [x MOVE IN MANNER α] (=x's act causes x to move in a manner specified by α)

- e. Non-volitional motion (*roll*-verbs): [x MOVE IN MANNER α] (=x moves spontaneously in a manner specified by α)

    f.   Accomplishment (e.g. *dry* tr.): [x ACT] CAUSE [BECOME [y BE AT-[$_{\text{STATE}}$ α]]]
           (=x's act causes y to become α)

    g.   State (e.g. *belong*): [y BE AT-[LOCATION OR STATE α]] (=y is at α)

LCS representations are composed of semantic predicates and their semantic arguments. The semantic role of an argument, such as agent or theme, can be read off from the semantic property of the predicate it occurs with. To illustrate, the meaning of the verb *dance* in (5a) is described as an act of the argument *x* engaging in a particular activity represented by ACT IN A DANCING MANNER. An activity (represented by ACT or ACT ON) denotes an action with no specified endpoint (i.e. 5a, b). On the other hand, an achievement (with change represented by BECOME) refers to a change of state that takes place instantaneously (i.e. (5c)). A motion event (MOVE) can be either volitional which is represented as being caused by the agent's volitional ACT (i.e. (5d)) or non-volitional, which is represented as a simple motion (i.e. (5e)). An accomplishment (represented by the schema "ACT CAUSE BECOME") refers to a complex event in which a change of state is caused by some activity (i.e. (5f)). A state (represented by BE) describes a circumstance that does not involve any change (i.e. (5g)).

    Given the schemas in (5), the morphological process of event noun formation is viewed as integration of a part or whole of the source verb's LCS into a single nominal concept of one of several aspectual types: Action, Event, Motion, Change, State, etc. The meaning of the deverbal noun *odori* 'dance', for example, will be represented in the (simplified) LCS, as in (6a), and that of the deverbal noun *bure* 'unstableness' as in (6b).

(6)   a.   The verb *odoru* 'to dance': [x ACT IN DANCING MANNER]
           → nominalized *odori* 'a dance': the act of [x ACT IN DANCING MANNER]
              i.e. *odori* is the act of x acting in a dancing manner.

      b.   The verb *bureru* 'to blur': [BECOME [y BE NOT-[FIXED AS EXPECTED]]]
           → nominalized *bure* 'a blur': the state of [y BE NOT-[FIXED AS EXPECTED]]
              i.e. *bure* is the state of y not being fixed as expected.

(6a) illustrates how the action noun *odori* 'dance' acquires its action meaning by taking over the whole of the source verb's LCS; (6b) shows that the state meaning of the noun *bure* 'a blur' obtains by picking out the state component [y BE NOT-FIXED AS EXPECTED] from the base verb's LCS [BECOME [y BE NOT-[FIXED AS EXPECTED]]], thereby *bure* signifies a certain state which is not stabilized in an expected way.

    Conversion also gives rise to nouns that denote concrete or abstract entities, and they can be classified into several subtypes, according to their meanings, as exemplified by (7).

(7) Entity nouns

    a.   Product, result: *kangae* 'thought' (< *kangae-ru* 'to think'), *tutumi* 'package' (< *tutum-u* 'to wrap'), *katamari* 'lump' (< *katamar-u* 'to stiffen'), *koori* 'ice' (< *koor-u* 'to freeze')

    b.   Agent : *suri* 'pickpocket' (< *sur-u* 'to pick someone's pocket'), *nagasi* 'strolling musician' (< *nagas-u* 'to play (music)')

    c.   Theme: *nagare* 'flow' (< *nagare-ru* 'to flow'), *tumami* 'snack' (< *tumam-u* 'to pinch, munch'), *turusi* 'ready-made clothes' (< *turus-u* 'to hang')

    d.   Instrument: *hakari* 'scale' (< *hakar-u* 'to measure'), *hataki* 'duster' (< *hatak-u* 'to dust'), *hurui* 'sieve' (< *huru(w)-u* 'to sieve'), *suki* 'plow' (< *suk-u* 'to plow')

    e.   Place: *toori* 'street' (< *toor-u* 'to pass'), *mawari* 'circumference' (< *mawar-u* 'to go around'), *sumai* 'dwelling' (< *suma(w)-u* 'to dwell')

    f.   Time: *owari* 'end' (< *owar-u* 'to end'), *hazimari* 'beginning' (< *hazimar-u* 'to begin')

Some of these examples will be attributed to creative formation and others to productive formation. If we can properly leave out creative coinages and limit our analysis to productive formation, then the meanings of derived nominals could plausibly be derived from the information contained in the base verbs. Kageyama (1999) and Ito and Sugioka (2002) develop an analysis where entity interpretations are derived by foregrounding (or focusing) a variable position in LCS. Abstracting away from technical details, the 'place' meaning of the noun *sumai* 'residence' (< *suma-u* 'to dwell') in (7e), for example, obtains as schematically shown in (8), where the locative argument (z) of the source verb's LCS is focused on to denote the place meaning of the noun *sumai*.

(8) The verb *sumau* 'to dwell': [[x ACT] CAUSE [x BE AT-z AS RESIDENT]]
    → Focusing on the place argument (z)
    the deverbal noun *sumai* 'residence': z such that [[x ACT] CAUSE [x BE AT-z AS RESIDENT]]
    i.e. the noun *sumai* means 'the place z where x lives'.

In a similar vein, the product or result interpretation as in (7a) will be derived by focusing on the theme argument (i.e. subject of BE), while the agent interpretation as in (7b) by focusing on the agent (i.e. subject of CAUSE, ACT or DO) in LCS.

    An interesting syntactic difference between event nouns like those in (4) and entity nouns like those in (7) manifests itself in the phenomenon called "argument inheritance", namely, transfer of the arguments of a source verb to a derived noun.

In general, the event nouns allow argument inheritance while the entity nouns do not. Observe the contrast between (9a) and (9b) discussed by Sugioka (1989).

(9) a. Event noun
 ringoku   *to*   *arasou* / *ringoku*   *to*   *no*   *arasoi*
 neighboring.country with fight   neighboring.country with GEN fight
 'to fight with a neighboring country'   'a fight with a neighboring country'

   b. Entity noun
 saihu   *o*   *suru* / **saihu*   *no*   *suri*
 wallet ACC pick   wallet   GEN   pickpocket
 'to steal a wallet'   lit. 'a pickpocket of a wallet'

In (9a), the deverbal noun *arasoi*, denoting a fighting event, takes the same argument ('with a neighboring country') as its source verb *araso-u* 'to fight'. This is a typical case of argument inheritance. In (9b), on the other hand, the deverbal noun *suri* designates a person, so it is unable to take a theme argument *saihu no* 'of a wallet' that corresponds to the direct object of the source verb *sur-u* 'pick'. Thus the general rule is that deverbal entity nouns are incapable of taking over the arguments of their source verbs. In Section 5, argument inheritance will be discussed in greater detail in connection with the formation of complex predicates involving deverbal compounds.

To sum up, the formation of deverbal nouns from infinitive verbs through conversion is subject to a number of phonological, semantic, and lexical restrictions, none of which is sufficient to cover all of the relevant phenomena. What makes the matter nebulous is the fact that such nouns are not accompanied by any particular morphological clue for predicting their specific meaning in systematic manner. Presumably, it is this intrinsic indeterminacy of semantic interpretations that enhances the speaker's linguistic creativity to exploit deverbal nouns in giving names to entities. In the next subsection, it will be observed that deverbal nouns gain more transparency and increased productivity when embedded in compound structure.

2.2 Deverbal nouns placed in syntactic and compound structure

A distinctive feature of converted nouns that has been commonly pointed out in the literature but does not have a satisfactory account is the fact that many of the putatively unacceptable deverbal nouns become fully acceptable if they are combined with another word in syntactic or morphological structure.

As a preliminary to this discussion, a peculiarity of deverbal nouns that has been generally overlooked in the literature is to be noted (Taro Kageyama, p.c.): certain deverbal nouns that are normally unable to stand alone become well-formed

when embedded in fixed, mostly idiomatic constructions, as shown in (10). (# indicates that the word in question is not normally used in isolation.)

(10) a. *#nige* (< *nige-ru* 'to flee, get away') → *nige o utu* [get away ACC attempt] 'back out, dodge questions/accusation'

b. *#nari* (< *nar-u* 'to sound') → *nari o hisome-ru* [sound ACC hush] 'be quiet and inactive'

c. *#kiri* 'end' (< *kir-u* 'to cut') → *kiri ga yoi* [end NOM good] 'be a good place to stop', *kiri ga nai* [end NOM non-existent] 'be endless'

d. *#heri* 'reduction' (< *her-u* 'to decrease') → *denti no heri* [battery GEN reduction] 'running down of a battery'

Given that the deverbal nouns in (10) are legitimate only in certain collocational combinations, these constructions can be regarded as a means of providing the deverbal nouns with support for their syntactic realization.

A more general means to support the deverbal bound nouns, which is often noted in the literature, is morphological compounding. Deverbal nouns that are totally or partially unacceptable as independent words become fully acceptable when they appear in compound structure. For example, although **ire* (< *ire-ru* 'to put in') is unacceptable in isolation, it is fully accepted if it appears as a constituent of compound words, as in *meisi-**ire*** [card-put.in] 'card case' and *hi-**ire*** [fire-put.in] 'igniting'. Under the assumptions of word-based morphology (Aronoff 1976), a larger word is built by putting together two or more freestanding words that are acceptable in themselves. The formation of compound words from non-freestanding deverbal nouns would thus present a different picture of morphology than the one postulated in Aronoff's (1976) theory if there is no principled way to explain this phenomenon.

In discussing compound words involving deverbal nouns, it is instructive to distinguish two types, depending on how compounding and conversion take place.

(11) Two types of deverbal compound nouns
Type I: A verb-verb compound verb is first produced and is nominalized as a whole to produce a compound noun or a compound verbal noun (VN).
$[V] + [V] \rightarrow [V\text{-}V]_V \rightarrow [[V\text{-}V]_V]_{N/VN}$
(e.g. *kasi-dasu* → *kasi-dasi*)

Type II: A simplex verb is first nominalized and then is compounded with another element to produce a compound noun or a compound VN.
$[V] \rightarrow [V]_{N/VN} \rightarrow [X \text{ - } [V] _{N/VN}]_{N/VN}$
(e.g. *dasi-ire* ⇍ **dasi-ireru*, *kane-moti* ⇍ **kane-motu*)

Type I concerns cases where a whole V-V compound is made first (cf. Chapter 8 [Kageyama, this volume]) and then undergoes conversion or nominalization, so that the compound takes the form of $[[V-V]_V]_N$ or $[[V-V]_V]_{VN}$. For example, the noun *kasi-dasi* 'checkout, loan' comes from the compound verb *kasi-dasu* [lend-send.out] 'to lend out'. Compounds of this type are parallel to the simplex event nouns in (4) in representing activities or events and are subject to argument inheritance. The theme object of the base verb *kasi-dasu* 'to lend out', for example, can manifest itself as a genitive phrase used with the deverbal compound *kasi-dasi* as in *hon no kasi-dasi* [book GEN lend-out] 'lending out a book'. Despite their event reading, many compounds of this type are not readily combined with the light verb *suru* 'do' to inflect for tense, as in ?**kasi-dasi-suru* [lend-out-do].

The failure to combine with *suru*, however, does not necessarily prove that the deverbal compounds of Type I are not VNs. This is because there is the possibility that a principle of lexical blocking is at work here to exclude the expected combination with *suru*. For example, the already existing V-V compound verb *kasi-dasu* [lend-out] pre-empts the creation of the more complex form ?**kasi-dasi-suru* [lend-out-do]. Additional examples of the same kind are given in (12), which shows these deverbal compounds are resistant to combination with *suru* because of the existence of V-V compound verbs.

(12) a. ?**kumi-tate-suru* [assemble-set.up-do] (*kumi-tate* [assemble-set.up] 'assembly' < *kumi-tate-ru* [assemble-set.up-PRS] 'to put together')

      b ?**uke-tuke-suru* [receive-attach-do] (*uke-tuke* [receive-attach] 'reception' < *uke-tuke-ru* [receive-attach-PRS] 'to accept, receive')

      c. **uti-kiri-suru* [strike-cut-do] (*uti-kiri* [strike-cut] 'closure' < *uti-kir-u* [strike-cut-PRS] 'to discontinue, cut short')

      d. ?**uri-dasi-suru* [sell-send.out-do] (*uri-dasi* [sell-send.out] 'promotion (of a new product)' < *uri-das-u* [sell-send.out-PRS] 'to market (a new product)')

Naturally, lexical blocking also applies to the combination of deverbal nouns based on simplex verbs and *suru*, as shown in (13) (Kageyama 1982).

(13) a. **hasiri-suru* [run-do] (*hasiri* [run] 'running' < *hasir-u* [run-PRS] 'to run')

      b. **kai-suru* [buy-do] (*kai* [buy] 'buying' < *ka-u* [buy-PRS] 'to buy')

The nouns *hasiri* 'running' and *kai* 'buying' are accepted as independent words but cannot form complex predicates with *suru*. Interestingly, they become compatible with *suru* when they are compounded as in (14).

(14)   a.   *[ko-basiri]-suru* ([short.step-run]-do) (< *ko-basiri* [short.step-run] 'run with short steps')   Cf. **ko-basir-u* [short.step-run-PRS]

       b.   *[syoodoo-gai]-suru* ([impulse-buy]-do) (< *syoodoo-gai* [impulse-buy] 'impulse buying')   Cf. **syoodoo-ga-u* [impulse-buy-PRS]

The deverbal compounds can be compounded with *suru* because they cannot inflect by themselves, as shown by the total ungrammaticality of **ko-basir-u* [short.step-run-PRS] and **syoodoo-ga-u* [impulse-buy-PRS].

   Much the same observation can be made with simplex deverbal nouns like a one-mora noun **ki* 'wearing' (< *ki-ru* 'to wear') and a two-mora noun **hore* 'falling in love' (< *hore-ru* 'to fall in love') which are unable to stand alone for phonological or other reasons. These simplex nouns cannot be directly followed by *suru*, either, as in **ki-suru* [wear-do] and **hore-suru* [fall.in.love-do]. Curiously, however, combinations with *suru* become entirely acceptable if the simplex nouns are turned into compounds, as in [*atu-gi*]-*suru* ([thick-wear]-do) 'dress heavily' and [*hitome-bore*]-*suru* ([one.glance-fall.in.love]-do) 'fall in love at first sight'.

   Returning to (11), Type II deverbal compounds can be conveniently divided into three classes in terms of their internal composition (cf. Nishio 1988: 58–59).

(15)   a. $[[V]_{VN}\text{-}[V]_{VN}]_{VN}$ in coordinate relation
            dasi-ire [put.in-take.out] 'putting in and taking out', *kasi-kari* [lend-borrow] 'borrowing and lending', *nori-ori* [get.on-get.off] 'getting on and off'

       b. $[[V]_{VN/N}\text{-}[V]_{VN}]_{VN}$ in modification relation
            kai-gui [buy-eat] 'spending one's pocket money on sweets', *kiri-uri* [cut-sell] 'selling by the piece', *tati-yomi* [stand-read] 'browsing while standing'

       c. $[[X]\text{-}[V]_N]_N$
            kane-moti [money-possess] 'rich man' (agent), *kubi-maki* [neck-wind.around] 'muffler' (instrument), *musi-sasare* [bug-bite.PASS] 'a bug bite' (product), *ato-mawasi* [afterward-pass] 'deferment' (activity), *hi-ire* [fire-put.in] 'kindling' (activity), *yuki-doke* [snow-melt] 'thawing of snow' (event), *yo-ake* [night-end] 'dawn' (time), *mono-oki* [thing-store] 'storeroom' (place)

The examples in (15a) involve coordinate or *dvanda* compounds, with two constituents arranged in 'and' or 'or' relations, whereas those in (15b) denote activities where the first deverbal noun modifies the second one in a variety of semantic relations. These two classes are characterized as VNs (verbal nouns) because they can combine with *suru* to represent tense. On the other hand, the examples in (15c), consisting of a noun or adverbial ([X]) and a deverbal noun, designate a diversity of concepts like a person, instrument, product, activity, and time. These compounds are characterized as simple Ns, whether entity nouns or event nouns; they are not VNs because they

fail to combine with *suru* to represent tense (e.g. *[ato-mawasi]-suru* ([afterward-passing]-do) 'defer').

The three classes of Type II deverbal compounds are thus analyzed in the morphological structures shown in (15a, b, c); the examples in (15a, b) are compound VNs while those in (15c) are compound nouns. Concerning the (15c)-type compounds, Ito and Sugioka (2002), in line with Di Sciullo and Williams' (1987) analysis of French synthetic compounds, contend that they originate from a syntactic verb phrase (V´) and are reanalyzed as a noun in an exocentric structure [[N-V]$_{V'}$]$_N$. Their analysis, however, evades the question of how the reanalysis takes place. In addition, provided that the V´ structure in syntax only consists of a head verb and its internal argument, it cannot cope with cases in which the first constituents do not correspond to internal arguments but to something else such as the passive agent in *musi-sasare* [bug-bite.PASS] 'bug bite' or the adverbial in *too-mawasi* [far-turn] 'roundabout'.

As mentioned above, the deverbal compounds in (15a, b, c) are problematic for the theory of word-based morphology because their individual constituents do not normally occur as independent words but can make up well-formed words only when compounded with other elements. Furthermore, their formation is highly productive, unlike the unproductive *cranberry* compounds in English, which Aronoff (1976) deals with in terms of morphological reanalysis. Increased productivity of conversion in compound structure can also be observed with deadjectival nouns that occur as the head (right-hand constituent) of a compound. Like deverbal nouns, deadjectival nouns become acceptable in compound structure even though they are totally unacceptable in isolation. In each of the compounds of (16), the second constituents are nouns originating from adjectives (to be precise, bare stems of adjectives).

(16) a. Deadjectival compound nouns
           asi-garu [foot-light] 'foot soldier' (< *karu-i* 'light'), *hara-ita* [stomach-painful] 'stomachache' (< *ita-i* 'painful'), *en-daka* [yen-high] 'strong yen rate' (< *taka-i* 'high'), *yo-naga* [night-long] 'long night' (< *naga-i* 'long')

      b. Deadjectival compound ANs
           haba-hiro [width-broad] 'broad' (< *hiro-i* 'wide'), *hara-guro* [stomach-black] 'malicious' (< *kuro-i* 'black'), *iro-ziro* [color-white] 'fair-skinned' (< *siro-i* 'white'), *kuti-garu* [mouth-loose] 'loose-tongued' (< *karu-i* 'light')

The examples in (16a), which represent concrete entities or states, are considered as simple nouns, while those in (16b), representing properties or attributes, would be analyzed as ANs (adjectival nouns) or noun-like adjectives because they are inflected with *na* when used prenominally (e.g. *iro-ziro na*) and with the copula *da* when used in the tensed position (e.g. *iro-ziro da*). In terms of the syntactic categories of a

whole compound, then, the conversion from verbs in (15) and the one from adjectives in (16) share an intriguing property that is not observed in languages like English. That is, the whole compounds either show up as simple Ns (15c, 16a) or take over the characteristic of verbs or adjectives to make up the hybrid categories of VNs or ANs (15a,b, 16b).

To recapitulate, V-to-N and A-to-N conversions are hardly productive when applied to simplex verbs or adjectives but gain increased productivity if the output is combined with another morpheme to make up a compound word or an idiomatic phrase (cf. (10)). Deverbal nouns have a wide range of meanings from events, states, or properties to concrete objects like agents, instruments, products, places, and time. By contrast, the semantic range of deadjectival nouns is limited to states and properties and rarely extends to concrete entities. This disparity could presumably be attributed to the difference in semantic properties and argument structures of verbs and adjectives, which requires further research to be defined.

3 Internal relations of deverbal compounds

This section discusses the semantic relations between the first constituents of deverbal compounds and their second constituents. Preliminary to the discussion, let us compare the compound and noun phrase structures that share the internal composition of "Adjective + Noun". For example, both the noun phrase *huru-i hon* [old-INFL book] 'an old book' and the compound noun *huru-hon* [old-book] 'a used book' appear to have a simple modification relation inside them. As to the noun phrase it is interpreted as a conjunction of two propositions "x is a book" and "x is old". Anything that is a book and is old is referred to as a *huru-i hon*. The compound, on the other hand, does not convey the same meaning: *huru-hon* is not characterized merely as "x is a book" and "x is old" but designates used or secondhand books, where *huru-* lit. 'old' serves to distinguish used books from brand-new books. Thus compound formation has a special property of creating a concept that is subordinate to the denotation of the head noun. This kind of interpretation can be referred to as "semantic specification", which is also relevant for deverbal compound formation. This section will discuss how semantic specification is implemented in deverbal nouns, using LCS and Qualia Structure as lexical semantic representations.

3.1 Semantic specification in root compounds

Before going into the details of deverbal compound nouns, we will briefly review previous analyses of Japanese root compounds of the form "X + N". It is well known that Japanese root compounds have a strong tendency to conform to the "*IS A* Condition" of Allen (1978) or the Right-hand Head Rule of Williams (1981), namely

the right-hand element in principle denotes the superordinate term of a whole compound. This type of compound can be classified into several types according to how the first noun (N1) specifies the denotation of the second noun (N2), as exemplified in (17). For more examples, see S. Yumoto (1978) and Kageyama (1993), among others.

(17) a. Attribute/mode: *boozu-atama* [bonze-head] 'a shaven head', *kiku-zara* [chrysanthemum-dish] 'a plate shaped like a chrysanthemum'

      b. Ingredient/material: *garasu-do* [glass-door] 'a door made of glass', *kami-bako* [paper-box] 'paper box'

      c. Purpose: *ama-gutu* [rain-shoes] 'shoes to wear on a rainy day', *tya-doogu* [tea-utensils] 'utensils for a Japanese tea ceremony'

      d. Place/time: *koobe-biihu* [Kobe-beef] 'Kobe beef', *haru-kaze* [spring-wind] 'spring breeze'

Working in the framework of early generative grammar, Okutsu (1975) and Makino (1976) attempted to describe the internal semantic relations by assuming that the compound nouns would be derived from underlying clausal structures involving relative clauses or adjuncts by transformation-like rules operating in the lexicon. For example, *haru-kaze* 'spring breeze' was analyzed as deriving from a relative clause structure like *haru ni huku kaze* [spring in blow wind] 'wind that blows in spring'. Such a transformational approach was soon abandoned in favor of the later lexicalist approach.

More recently, researchers have begun to refine the lexical approach to compound nouns by utilizing the idea of Pustejovsky's (1995) Generative Lexicon theory, where not only the standardly assumed kind of lexical meaning but also what is generally regarded as encyclopedic and ontological knowledge is included in the lexical information of words. According to Pustejovsky (1995: 86), such lexical information can be described by the lexical semantic representation called "Qualia Structure" which includes the following four "roles": Formal role (that which distinguishes a given object within a larger domain such as orientation, magnitude, shape, dimensionality, color, and position), Constitutive role (the relation between the object and its constituents or proper parts, such as material, weight, parts, and component elements), Telic role (the purpose and function of the object; the purpose that an agent has in performing an act, built-in function or aim which specifies certain activities), and Agentive role (factors involved in the object's origin or things that bring it about, such as creator, natural force, or causal chain).

Following the classic model of Pustejovsky (1995), the description of the four roles is illustrated with the cooking utensil *nabe* 'pan, pot' in a simplified form, as in (18) (see Pustejovsky et al. (eds.) (2013) for recent developments of the theory).

(18) Qualia of *nabe* 'pan/pot'

$$
\begin{bmatrix}
\text{FORMAL ROLE: x: phys artifact} \\
\text{CONSTITUTIVE ROLE: part of (w, x)} \quad \text{w: container, handle, lid} \\
\text{TELIC ROLE: } \mathbf{cook}'(e_1, y, z, \text{with } x) \quad \text{z: food} \\
\text{AGENTIVE ROLE: } \mathbf{create}'(e_2, i, x)
\end{bmatrix}
$$

In (18), the Formal role represents *nabe* as a physical object and an artifact; the Constitutive role provides information about the internal parts that make up a typical *nabe* (e.g. handle, container, etc.); the Telic role specifies the stereotypical purpose for which a pot is intended (i.e. it is intended for use in cooking food); and finally, the Agentive role describes how the object comes about (i.e. it is created by a person).

Johnston and Busa (1996) exploit these four qualia representations to account for the representative relations of modification in root compounds. Specifically, they argue that the modifier constituent of a compound makes the meaning of the head constituent specific by picking out suitable information from one of the four roles. The gist of Johnston and Busa's analysis is that in Italian, characteristic prepositions that appear inside compounds provide morphological clues to identify which role is picked out. To be concrete, in *coltello da pane* [knife for bread] 'bread knife', the preposition *da* 'for' specifies the purpose or use of the knife (Telic role), and in *succo di limone* [juice from lemon], the preposition *di* 'from' points to the source or ingredient of the juice (Agentive role).

Kageyama (1999) applied Johnston and Busa's Qualia Structure analysis to Japanese root compounds (where neither postpositional nor prepositional markers show up). By way of an illustration, we will examine in (19) below how the compound nouns taking the form of "X-*nabe*" (X-pot) are interpreted by referring to one of the four roles of Qualia Structure.

(19) a. The modifier *X* specifies an external characteristic (i.e. Formal role) of the head noun.
        maru*-*nabe [round-pan]
        FORMAL ROLE: x: phys artifact, x: **round**
        i.e. ***maru*-*nabe*** is a pan that is round in shape.

    b. The modifier *X* specifies an internal characteristic (i.e. Constitutive role) of the head noun.
        ryoote*-*nabe [both.hands-pan]
        CONSTITUTIVE ROLE: part of (h, x), h: **two handles**
        i.e. ***ryoote*-*nabe*** is a pan that has two handles.

    c.   The modifier *X* specifies a stereotypical or intended purpose (i.e. Telic role) of the head noun.
        ***tenpura**-nabe* [tempura-pan]
        TELIC ROLE: **cook**´$(e_1, y, z,$ with $x), z:$ ***tempura***
        i.e. ***tenpura**-nabe* is a pan for cooking tempura.

    d.   The modifier *X* specifies the origin (i.e. Agentive role) of the head noun.
        ***Nanbu**-nabe* [Nanbu (place name)-pan]
        AGENTIVE ROLE: **create**´$(e_2, i, x,$ in $w), w:$ ***Nanbu***
        i.e. ***Nanbu**-nabe* is a pan made in Nanbu.

As seen from the illustrated formulas, the lexical meaning of the left-hand constituent is inserted into a suitable position of one of the four roles (i.e. *maru-* 'round' in the Formal role, *ryoote-* 'two handles' in the Constitutive role, *tenpura-* in the Telic role, and *nanbu-* in the Agent role), resulting in a specification of the denotation of the head noun. It is an open question, however, whether the four roles can exhaust all the possible interpretations of compound words. Consider, for example, the compound noun *bunka-nabe* [culture-pot], which refers to a heavy double-handled pot used to cook rice (Taro Kageyama, p.c.). This type of pot was invented in the mid-1900s, when many new inventions were named by compounds with the modifier *bunka* 'cultured or advanced' to stress the modernity. Due to its naming function, *bunka* would probably have no place to fit in with any of the four qualia representations of *nabe* 'pot'.

3.2 Semantic specification in deverbal compound nouns

We will now turn to the central topic of this chapter, i.e. compound words headed by deverbal nouns, and discuss the internal relations between the two constituents of a compound in terms of argument and adjunct functions of the first constituent in relation to the second. In the generative approach to English morphology, it is widely acknowledged that the internal relations of deverbal (or synthetic) compounds like *bird-watching* and *windsurfing* are constrained by the following two conditions (cf. Lieber 1983).

(20)   a.   The first constituent (N) is interpreted as the internal argument of the second constituent (V) if it has one, as in *bird-watching*,

       b.   or otherwise as a semantic argument (adjunct) if the verb lacks an obligatory internal argument, as in *windsurfing*.

Selkirk (1982) proposes a similar condition stipulating that all "non-subject" arguments of a lexical category *X* must be satisfied within the first order projection (i.e.

as the immediate sister) of *X*. Lieber's and Selkirk's proposals are essentially the same except that Lieber employs the notion "internal argument" (i.e. argument that appears in the sister position of a verb in syntactic structure) while Selkirk refers to "non-subject" arguments. Japanese data supports Lieber's condition. Selkirk's is too strong because it incorrectly rules out the subjects of unaccusative verbs (see Kageyama 1982, 1993). Consider the possible and impossible combinations summarized in (21).

(21)   [N-[V]$_{N/VN}$] compounds

    a.   Object + transitive verb: *sakana-turi* [fish-angle] 'angling', *moti-tuki* [rice.cake-pound] 'rice-cake making'

    b.   Complement + intransitive verb: *ie-de* [home-go.out] 'running away from home', *inaka-zumai* [country-live] 'country life'

    c.   Subject + unaccusative verb: *zi-nari* [earth-rumble] 'underground rumbling', *me-mai* [eye-swirl] 'dizziness'

    d.   *Subject + transitive verb: *sirooto-zukuri [amateur-make] 'made by an amateur', *titi-sodate [father-raise] 'raised by one's father'

    e.   *Subject + unergative verb: *kodomo-asobi [kid-play] 'child's playing', *onna-bataraki [woman-work] 'women's working'

In Japanese, three types of arguments as in (21a, b, c) qualify as the internal arguments of verbs stipulated in Lieber's first condition in (20a), whereas the subjects of transitive verbs (21d) and unergative verbs (21e) cannot be accepted. Since the notion of internal argument pertains to "argument structure", we can safely assume that compounds given in (21a, b, c) are produced by inserting a particular noun in the open slot for the internal argument of the deverbal noun. The schema in (22) exemplifies how the noun *sakana* 'fish' is interpreted as the internal argument (*y*) of the deverbal noun *turi* 'angling'.

(22)   argument structure of *turi* 'angling':   x <y>,   y: **sakana** 'fish'
     i.e. **sakana**-*turi* means 'angling for fish'

    The Japanese data in (21a, b, c) appear to confirm the overall validity of not only Lieber's (1983) condition but also the Unaccusative Hypothesis (i.e. the subject of an unaccusative verb occupies the same syntactic position as the object of a transitive verb). There are only a few exceptions to the internal argument restriction, as Kageyama (1982) observes with examples like those in (23).

(23) a. Indirect object: *hito-makase* [person-leave] 'leave one's job to other people'

   b. Transitive subject: *oya-yuzuri* [parent-hand.over] 'inherited from one's father and/or mother'

What distinguishes the examples in (23) from those in (21) is the presence or absence of "event" interpretation. The deverbal compounds in (21a, b, c) denote events and can occur as the object (marked by *o*) or subject (marked by *ga*) of the verb *suru* 'do' as in (24).

(24) a. *usagi-gari o suru* [rabbit-hunt ACC do] 'to do rabbit hunting'

   b. *ie-de o suru* [home-go.out ACC do] 'to run away from home'

   c. *me-mai ga suru* [eye-swirl NOM do] 'feel dizzy'

The examples in (23), by contrast, denote states or attributes and do not occur in the syntactic frame "X *ga/o suru*". This issue is taken up in some detail in Section 4.

   We will now turn to the second condition of Lieber (1983), given earlier in (20b): the first constituent (N) of a deverbal compound is interpreted as a semantic argument (adjunct) if the second constituent (V) lacks an obligatory internal argument. The initial part of (20b) (i.e. the availability of an adjunct interpretation) holds true for Japanese, although the following statement indicated by the "if" clause does not. Here we focus on the initial statement of (20b) on adjunct interpretation, reserving discussion on the subsequent condition until Section 5. Consider examples like those in (25), where the first constituents are adjectival bare stems.

(25) [A-[V]$_{N/VN}$] compounds
   a. Adjectival stems describe manners of the action denoted by deverbal nouns: *haya-aruki* [fast-walk] 'fast walking' (< *aruk-u* 'to walk'), *yasu-ukeai* [ready-undertake] 'making a promise too readily' (< *ukea-u* 'to undertake'), *taka-warai* [high-laugh] 'loud laughter' (< *wara-u* 'to laugh')

   b. Adjectival stems describe modes of the change denoted by deverbal nouns: *oso-zaki* [late-bloom] 'late bloom' (< *sak-u* 'to bloom'), *naga-wazurai* [long-be.ill] 'long disease' (< *wazura(w)-u* 'to be ill')

   c. Adjectival stems describe resultant states of the events denoted by deverbal nouns: *taka-domari* [high-stop] '(said of stock exchange) being stuck at a high level' (< *tomar-u* 'to stop'), *siro-nuri* [white-paint] 'being painted white' (< *nur-u* 'smear')

A prominent feature of these compounds is that despite the superficial appearance of "A+N", the left-hand constituents (i.e. adjectival stems) do not directly modify

the nouns in the same way as an adjective modifies a head noun in noun phrase structure. Rather, the interpretation of adjectival stems should be derived from the semantics of the verb the right-hand noun is based on. For example, *taka-warai* [high-laugh] 'horselaugh, boisterous laugh' in (25a) does not mean 'a high/loud laughter' in which the adjective directly modifies the noun *warai*, but 'laughing loudly or boisterously', where *taka-* is interpreted as describing the manner of 'laughing'. Since the bare stems of adjectives (e.g. *taka-*) have the potential to function as adverbs with the addition of *-ku* inflection, it is perfectly natural that the adjective stems in (25) are construed as "adverbial". It is also natural that the nouns in the right-hand position accept such an adverbial interpretation because they are formed based on the verbs.

To be more specific, we can assume, following Ito and Sugioka (2002), that the deverbal nouns in the head position "inherit" the lexical semantic representations (or LCS) of their source verbs. For example in *haya-aruki* [fast-walk], *aruki* inherits the semantic structure of the motion verb *aruku* 'walk', and the modifier *haya-* 'fast' supplies the meaning 'fast' to specify its manner component, giving rise to the interpretation of 'fast walking'. Likewise in *siro-nuri* [white-paint], the modifier *siro-* 'white' specifies the resultant state ('be painted') of the painting event. In fact, these compounds with the superficial combination of an adjectival stem and a nominalized verb can be plausibly paraphrased as *haya-ku aruku* 'walk quickly' and *siro-ku nuru* 'paint white' with the *-ku* inflection attached to the adjectival bare stems. The semantic structures of *haya-aruki* and *siro-nuri* will thus be schematically represented as in (26). The semantic representation in (26a) is based on the LCS schema of verbs of volitional motion (cf. (5d) above), and that in (26b) on the LCS schema of accomplishment verbs (cf. (5f) above).

(26)   a.   *haya-aruki* 'fast walking'
            [x ACT] CAUSE [x MOVE IN MANNER α], α: **fast**
            i.e. *haya-aruki* is the act of walking fast.

       b.   *siro-nuri* 'painted white'
            [x ACT] CAUSE [BECOME [y BE [AT-COLOR α]]], α: **white**
            i.e. *siro-nuri* is the state of being painted white.

In (26a), the modifier *haya-* 'fast' enriches the source verb's meaning with a specific manner of motion; in (26b), the modifier *siro-* 'white' does so with a specific resultant state so that the resultant state of the action represented by *nuri* is foregrounded. These modifiers thus serve to specify an underspecified part of the head verb's LCS.

This approach can be extended to an analysis of deverbal compound nouns whose first constituents are nouns representing adjuncts. This type of compound is far more productive than the adjectival type in (25). Typical examples are shown in (27).

(27) [N-[V]$_{N/VN}$] compounds

    a. Modifiers specify manners of the action or change denoted by deverbal head nouns.
       inu-kaki [dog-paddle] 'dog paddle' (< *kak-u* 'to paddle'), *inu-zini* [dog-die] 'a useless death' (< *sin-u* 'to die')

    b. Modifiers specify instruments employed in the events of deverbal head nouns.
       mizu-arai [water-wash] 'washing in water without using soap' (< *ara-u* 'to wash'), *pen-gaki* [pen-write] 'writing with a pen' (< *kak-u* 'to write'), *ami-yaki* [gridiron-grill] 'grilling on a grid' (< *yak-u* 'to grill')

    c. Modifiers specify resultant states of the events denoted by deverbal head nouns.
       wa-giri [circle-cut] 'round slices' (< *kir-u* 'to slice'), *yama-zumi* [mountain-load] 'a large pile' (< *tum-u* 'to load')

In (27a), *inu* 'dog' in *inu-kaki* [dog-paddle] is construed as a metaphoric representation of the manner of a dog's swimming, in (27b), *mizu* 'water' in *mizu-arai* [water-wash] describes the means used for washing, and in (27c), *wa* 'circle' in *wa-giri* [circle-cut] specifies the resultant state of the object which is cut so that the compound represents the state rather than the action. The schemas in (28) illustrate how these semantic interpretations obtain from LCS representations.

(28)   a. *inu-kaki* [dog-paddle] 'dog paddle'
       [x ACT IN MANNER α] CAUSE [x MOVE in water], α: **like a dog**
       i.e. *inu-kaki* is the act of swimming in a manner typical of a dog.

    b. *mizu-arai* [water-wash] 'wash in water'
       [x ACT ON y BY MEANS OF α], α: **water**
       i.e. *mizu-arai* is the act of washing something by means of water

    c. *wa-giri* [ring-cut] 'cut in round slices'
       [x ACT] CAUSE [BECOME [y BE [AT-SHAPE α]]], α: **circle**
       i.e. *wa-giri* is the act of cutting something into round slices

As seen in the representations in (28), adjunct interpretation is derived by inserting the concept designated by the left-hand element in a supplementary position (α) of the lexical semantic representation of a source verb to make it more specific.

    We have observed that Lieber's (1983) condition on English deverbal compounds generally holds for Japanese deverbal compounds as well. The first part of her condition on the internal argument relation, illustrated with the examples in (21a, b, c),

is formally captured by specifying the internal argument of the deverbal noun with the nominal concept given by the first constituent of a compound. The second part of Lieber's condition, namely, adjunct interpretation, on the other hand, will be formally represented by filling in the modifier position which is unspecified in the lexical semantic representation of the source verb. In both cases, semantic specification serves to create subordinate terms from superordinate terms by narrowing down the range of denotations.

4 External semantics of deverbal compounds

This section will survey the external semantics of deverbal compounds, namely, the denotations of deverbal compounds as a whole. The distinction between event nouns and entity nouns, discussed in Section 2.1 with respect to simplex words, is relevant to compound words as well, and the deverbal event nouns of the compound type are subject to essentially the same line of analysis as those of the simplex type. Representative types of event nouns are exemplified in (29).

(29) Deverbal event nouns
    a. Natural happening
        gake-kuzure [cliff-collapse] 'landslide'

    b. Physiological happening
        me-mai [eye-whirl] 'dizziness, the whirl of the brain'

    c. Change of state
        huna-yoi [ship-get.sick] 'get seasick'

    d. Action in a specific manner
        hitori-aruki [one.person-walk] 'walk by oneself'

    e. Action by means of a specific instrument
        pen-gaki [pen-write] 'write with a pen'

    f. Temporary state
        apaato-zumai [apartment-live] 'living in an apartment'

    g. Resultant state
        iro-tuki [color-attach] 'colored'

The meaning of an entire deverbal compound can be derived by turning the source verb's meaning into a nominal concept by filling a variable position or unspecified slot in its LCS with the semantic notion denoted by the first constituent of a compound as is exemplified in (30).

(30) a. LCS of the verb *kuzureru* 'collapse': [BECOME [y BE AT-COLLAPSED]]
        → the deverbal nominal *gake-kuzure* [cliff-collapse] 'landslide'
         the event of [BECOME [y BE AT-COLLAPSED]], y: **cliff**
         i.e. *gake-kuzure* is the event of a cliff's becoming collapsed.

     b. LCS of the verb *aruku*: [[x ACT] CAUSE [x MOVE]]
        → *hitori-aruki* [one.person-walk]
         the act of [[x ACT IN MANNER α] CAUSE [x MOVE]], α: **by oneself**
         i.e. *hitori-aruki* is the act of walking by oneself.

     c. LCS of the verb *sumau*: [[x ACT] CAUSE [x BE AT-z]]
        → *apaato-zumai* [apartment-live]
         the act of [[x ACT] CAUSE [x BE AT-z]], z: **apartment**
         i.e. *apaato-zumai* is the act of living in an apartment.

In parallel to the nominalization of simplex verbs exemplified earlier in (6), the source verbs in (30) are first turned into nominal concepts by conceptualizing their whole event or action as an integrated notion. The only way the compound nouns in (30) differ from the simplex nouns in (6) is that a variable in the semantic structure of the source verb is specified by the concrete notion designated by the left-hand element of the compound. Thus *gake-kuzure* in (30a) means not just a collapsing event but the event of a cliff's collapsing, and *hitori-aruki* in (30b) means not just the act of walking but the act of walking alone.

While the event nouns of type (29) are formed productively and regularly, deverbal compounds that function as entity nouns are created with less regularity. At the extreme end of irregularity are exocentric compounds for naming objects like those in (31) (Kageyama 2009; cf. also Chapter 6 [Namiki and Kageyama, this volume]).

(31) *saru-suberi* [monkey-slip.down] 'crape myrtle (tree)', *neko-irazu* [cat-need.not] 'rat poison', *oya-sirazu* [parent-know.not] 'wisdom tooth'

If such irregular coinage of exocentric compounds by naming is excluded, the possible range of meanings available for deverbal compounds is limited to several types of concrete notions shown in (32) (cf. Sugioka 2001; Kageyama 2009).

(32) Deverbal entity nouns
     a. Agent/person: *kane-moti* [money-have] 'rich man', *hana-uri* [flower-sell] 'flower vendor', *ato-tori* [footstep-take] 'successor', *mahoo-tukai* [magic-use] 'witch, wizard', *hora-huki* [brag-make] 'braggart'

     b. Instrument/means: *nezumi-tori* [mouse-catch] 'mouse catcher', *tume-kiri* [nail-cut] 'nail clipper', *te-huki* [hand-wipe] 'hand towel'

c. Place: *kuzu-ire* [trash-put.in] 'wastebasket', *kuruma-yose* [car-move.to] 'carport on the entrance of a house', *te-arai* [hand-wash] 'toilet'

d. Product: *yasai-itame* [vegetable-fry] 'fried vegetable', *tamago-yaki* [egg-fry] 'Japanese omelet', *mizu-tamari* [water-gather] 'puddle', *hada-are* [skin-get.rough] 'state of one's skin being rough'

e. Time: *yo-ake* [night-end] 'dawn', *tuyu-ake* [rainy.season-end] 'the end of a rainy season'

A few comments are in order concerning the details of the examples in (32). Agent nouns (32a) are actually ambiguous between a real entity (i.e. a person) and an attribute or state characterizing that person. Thus, *kane-moti* may refer to a particular person, as in *hitori no **kane-moti*** 'one rich man', or a temporary or permanent state of having a lot of money as in ***kane-moti** ni nari-tai* '(I) want to be rich'. Analogous examples are *saisi-moti* [wife.child-have] meaning 'the state of having a wife and children' or 'a man who has this attribute' and *byooki-moti* [illness-have] meaning 'the state of suffering from a long illness' or 'a person with that characteristic'. Instrument nouns (32b) refer to particular tools or devices, though some of them could be considered as being derived by clipping from longer words that explicitly represent the semantic category of the word, as in *nezumi-tori* [mouse-catch] 'mouse catcher' from *nezumi-tori-ki* [mouse- catch-device]. The availability of place nouns in (32c) will plausibly be attributed to the verb's ability to take a locative argument (cf. Y. Yumoto 2015), though there are also place nouns like *te-arai* [hand-wash] 'bathroom, toilet' where the base verb *arau* 'wash' does not select a locative argument. The formation of product nouns like those in (32d) takes place fairly productively. Time nouns (32e) might also be analyzed as a kind of product or result nouns in the sense that the time in question (e.g. *yo-ake* 'dawn') comes about as a result of the night's being over (cf. Kageyama 1993; Y. Yumoto 2015).

The diverse entity meanings of the compound nouns in (32) will obtain through the same mechanism that is responsible for the entity interpretations of converted simplex nouns exemplified earlier in (8) (Section 2.1). That is, an argument in the LCS representation of a source verb is focused on to represent that entity. For example, in *kane-moti* [money-have] 'rich man' (32a), the agent argument is focused on. The only difference is that the compound noun has another argument filled in with the left-hand element so that the range of the arguments that can be focused on is limited. As a result, *kane-moti* means a person who possesses (lots of) money. Likewise, in *kuzu-ire* [trash-put.in] 'wastebasket' (32c), the locative argument of the source verb *ireru* 'put in' is focused on, so that the whole compound denotes a receptacle to put trash in. Which argument in the lexical semantic representation is focused on, however, will be attributed to the need for naming or pragmatic context. There will be no logical reason why *nezumi-tori* [mouse-catch] 'a mouse catcher'

denotes an instrument while *syakkin-tori* [debt-collect] 'a debt collector' denotes a person, though they are based on the same source verb *toru*.

From a cognitive point of view, it might be thought that the deverbal compounds that function as entity nouns are all derived from their corresponding event compounds by metonymy. An agent, for example, can be conceived of as a metonymic representation of a whole act or event in which he/she is directly involved. This approach, however, is not easily compatible with the peculiarity of *rendaku* or sequential voicing in these compounds. The noun *tukai* 'use', for example, undergoes voicing in the event noun *kane-zukai* [money-spend] 'the way one spends one's money', but not in the agent noun *mahoo-tukai* (cf. **mahoo-zukai*) 'wizard'. In a similar vein, the noun *huki* 'wipe' undergoes voicing in the event noun *kara-buki* [dry-wipe] 'wipe with a dry cloth', but not in the means noun *te-huki* 'hand towel'. This evidence suggests that event nominalization and entity nominalization are separate processes (cf. Sugioka 2002).

As to the diversity of meanings in (32), one might attribute it to the sparseness of native suffixes. Although there are some suffixes that explicitly represent the semantic categories of derived nouns such as agentive suffixes (*-ka, -sya, -nin*: Chapter 17 [Ono, this volume]) or suffixes indicating places (*-ba*, *-syo*), most of them are of Chinese origin and hence are incompatible with native compound words. This argument, however, will not go through because instrument/means compounds like those in (32b) are fairly productive despite the existence of specific Sino-Japanese suffixes denoting a device (*-ki* 器), a machine (*-ki* 機), and the like that may be attached to native words. It is thus reasonable to assume that a special factor is at work in Japanese that promotes the formation of deverbal compounds that function as entity nouns as a means to enrich the native vocabulary.

Before closing this section, we will briefly touch on deadjectival compound nouns whose right-hand constituents are nominals originating from adjectival bare stems. Since there are few Japanese adjectives that take complements, most deadjectival compounds consist of a bare adjective and its subject, as exemplified in (33). For example in *iro-ziro* [complexion-white], *iro* 'complexion' corresponds to the (small) subject of the head *-ziro* 'white' (voiced from *siro*) in the clause *Hanako wa iro ga siroi* [Hanako TOP complexion NOM white], and the compound as a whole represents the fairness of the skin of a person embodied as the major subject *Hanako*.

(33) a. State or attribute: *iro-ziro* [complexion-white] 'fair-skinned', *ki-yowa* [mind-weak] 'timidity', *mi-omo* [body-heavy] 'pregnancy'

   b. Entity: *hoo-ziro* [cheek-white] 'meadow bunting', *me-ziro* [eye-white] 'Japanese white-eye (bird)', *o-naga* [tail-long] 'azure-winged magpie'

Obviously, the entity nouns in (33b) are due to creative coinage for naming rather than to a productive word formation rule.

In this section we have examined the external semantics of deverbal and deadjectival compound nouns. Except for naming in specialized or professional fields, the denotations of such compounds can be accounted for in a large measure by reference to the lexical semantic representations of the source verbs/adjectives. Based on the distinction of entity nouns and event nouns, we will move on to their syntactic properties in the next section.

5 Syntactic status of deverbal compounds

The present section will shift attention to the syntactic properties of deverbal compounds and explore how the semantic properties discussed in the foregoing section are reflected in their syntactic functions. Of particular importance is the distinction between simple nouns which serve only as referential expressions, and VNs (verbal nouns) which can function as predicates of clauses.

In her study of the syntax and semantics of English nominals, Grimshaw (1990) proposed to distinguish two classes of event nominals, in addition to the class of ordinary nouns denoting concrete or abstract entities ("result nominals" in Grimshaw's terminology): (i) "complex event nominals" or deverbal nouns that inherit the argument structure of the source verbs, and (ii) "simple event nominals" or nouns (not necessarily deverbal) which represent events but do not have argument structure. Drawing on Grimshaw's idea, Kageyama (1993) proposes a three-way distinction for deverbal compound nouns, as illustrated with some examples in (34).

(34)   a.   Entity nominal (Grimshaw's result nominal)
            mahoo-tukai 'wizard', *tume-kiri* 'nail clipper', *mono-oki* 'storeroom',
            mizu-tamari 'puddle', and other examples like those in (32).

       b.   Simple event nominal
            hosi-uranai [star-tell.fortune] 'horoscope', *hito-makase* [other.person-leave] 'leaving one's task to others', *otibo-hiroi* [fallen.ears.of.grain-pick.up] 'gleaning', *sara-arai* [dish-wash] 'dish washing', *me-mai* [eye-whirl] 'dizziness', *zi-hibiki* [earth-rumble] 'earth tremor', *kata-kori* [shoulder-get.stiff] 'the state of one's neck being stiff'

       c.   Complex event nominal (=Verbal Noun)
            ne-age [price-raise] 'to raise the price', *saki-okuri* [future-send] 'to put off', *sita-uti* [tongue-click] 'to click one's tongue', *yasu-uri* [cheap-sell] 'to sell cheap', *huna-yoi* [ship-get.sick] 'to get seasick', *naname-yomi* [slant-reading] 'to browse', *kara-buki* [dry-wipe] 'wipe with a dry cloth'

What concerns us here is the distinction between simple event nominals (34b) and complex event nominals (34c). Kageyama (1993: 187) points out that these two

classes of nominals exhibit distinct syntactic behaviors, despite their semantic similarity. The distinction lies in whether a given nominal can be directly followed by the light verb *suru* without mediation of case particles. The nouns exemplified in (34b) do not allow direct attachment of *suru* 'do' but instead must be accompanied with case particles like *ga* (nominative), *o* (accusative), or *ni* (dative) when they are used with *suru*. Observe the examples in (35), where the notation *(X)* means that the given expression will be ungrammatical if *X* is omitted.

(35) a. *Watasi   wa   totuzen   [**me-mai**]   *(ga)   si-ta.*
         I   TOP   suddenly   eye-whirl   *(NOM)   do-PST
         'I suddenly felt dizzy.'

   b. *Totuzen   [**zi-hibiki**]   *(ga)   si-ta.*
         suddenly   earth-rumble   *(NOM)   do-PST.
         'Suddenly the ground rumbled with a roar.'

   c. *Kyoo   wa   Ken   ga   [**sara-arai**]   *(o)   suru   ban   da.*
         today   TOP   Ken   NOM   dish-wash   *(ACC)   do   turn   COP
         'It's Ken's turn to do the dishes today.'

   d. *Watasi   wa   kessite   sigoto   o   [**hito-makase**]   *(ni)   si-nai.*
         I   TOP   never   job   ACC   others-leave   *(DAT)   do-NEG
         'I never leave my job to other people.'

The compounds exemplified in (34c), on the other hand, can be directly followed by (i.e. combined with) *suru*.

(36) a. *Zyookyaku   wa   zen'in   [**huna-yoi**]-si-ta.*
         passenger   TOP   all   ship-get.sick-do-PST
         'All the passengers got seasick.'

   b. *Kirin   ga   biiru   o   [**ne-age**]-si-ta.*
         Kirin (beer company)   NOM   beer   ACC   price-raise-do-PST
         'Kirin raised the price of beer.'

   c. *Ano   mise   wa   kitte   o   [**yasu-uri**]-si-te   iru.*
         that   shop   TOP   stamp   ACC   cheap-sell-do-GER   is
         'That shop sells stamps cheap.'

In Kageyama (1993), VNs are defined as those nouns that can be directly followed by *suru* to make a complex predicate in the so-called "light verb construction" (cf. Chapter 12 [Miyamoto and Kishimoto, this volume]). While Sino-Japanese words are prototypical VNs, the native compound nouns in (34c) are also qualified as VNs. The

compounds in (34b), on the other hand, do not belong to the category of VN in this regard because they do not occur in the light verb construction.

In addition, Kageyama (1993) observes that VNs share case-marking ability with genuine verbs and are capable of assigning case to their arguments even in clauses without tense, as illustrated by (37a) in which *sippitu* is a VN.

(37) a. *Sensei   ga   ronbun   o   sippitu   no   sai, . . .*
         professor   NOM   paper   ACC   write   GEN   time
         'When the professor was writing a paper, . . .'

   b. *Kirin   ga   biiru   o   oohabani   [**ne-age**]   no   sai,. . .*
       Kirin   NOM   beer   ACC   radically   price-raise   GEN   time
       'When Kirin raised the price of beer substantially, . . .'

In (37b), the native deverbal compound *ne-age* [price-raise] marks the object 'beer' with the accusative case in the same way as the Sino-Japanese VN *sippitu* 'to write' in (37a). Direct compatibility with *suru* and case-marking ability thus legitimately identify complex event nominals like those in (34c) as VNs.

The observations above lead us to modify (15c) and postulate three kinds of morphological structure for deverbal compound nouns in the form of [N-V], as shown in (38).

(38) a. Entity nouns (Grimshaw's result nominals) (=N)
        $[N - [V]_N]_N$, where the whole compound denotes an entity.

   b. Simple event nominal (=N)
      $[N - [V]_N]_N$, where the whole compound denotes an event, action, or state.

   c. Complex event nominal (=VN)
      $[N - [V]_{VN}]_{VN}$, where the whole compound denotes an event, action, or state.

(38b) and (38c) differ in the lexical category of a whole compound, N or VN, and this categorial distinction is assumed to be already available at the stage where a simplex verb undergoes conversion (nominalization).

Sugioka (2001, 2002) proposes to attribute this distinction between simple event nominals and complex event nominals (=VNs) to the relations of two constituents, generalizing that those deverbal compounds whose left-hand element serves as the argument of the source verb in the right-hand position are simple event nominals whereas only those whose left-hand element is an adverbial modifier of the source verb are complex event nominals that can be directly combined with *suru*. Actually,

however, complex event nominals allow arguments as well as adjuncts, as shown by *ne-age* [price-raise] 'to raise the price' and other examples (cf. Y. Yumoto 2010).

Now, recall the conditions on English deverbal compounds proposed by Lieber (1983) and Selkirk (1982), according to which the internal argument of the head derived from a verb, if there is one, must be realized within the compound structure. This condition fails to apply to complex event nominals of (34c)-type. Crucially, transitive verbs may be compounded not only with their objects, as in *ne-age(suru)* [price-raise (do)] 'to raise the price' and *sita-uti(suru)* [tongue-click (do)] 'to click one's tongue' in conformity with Selkirk's and Lieber's constraints, but also with a variety of adjuncts, as in *naname-yomi(suru)* [slant-read (do)] 'to browse' and *kara-buki(suru)* [dry-wipe (do)] 'wipe with a dry cloth'.

Why are the compounds of (34c)-type exempt from Lieber's and Selkirk's constraints? The reason is that these compounds function as predicates by themselves. Recall that transitive VNs of (34c)-type are able to take their own objects, as in *tukue o kara-buki-suru* [desk ACC dry-wipe-do] 'to wipe a desk with a dry cloth'. More generally, Japanese transitive verbs that are compounded with adjunct elements are allowed to realize their object arguments outside the compound structure. For example, the native compound verb *te-watasu* [hand-give] 'to pass by hand' consisting of an adjunct 'hand' and a ditransitive verb 'give' can take an accusative object and a dative object, as in *onnanoko ni tegami o te-watasu* [girl DAT letter ACC hand-give] 'to hand a letter to the girl' (cf. Chapter 7 [Kageyama, this volume]). The transitive compound VNs in (34c) such as *yasu-uri* [cheap-sell] 'to sell cheap' and *saki-okuri* [future-send] 'to put off' behave similarly to these compound verbs and can realize their object in the accusative case as can be observed in the examples of (36). By contrast, simple event nominals like those in (34b), which do not function as predicates, are not allowed to express their objects or other complements externally. They are thus similar to English deverbal compounds like *windsurfing*, which are nouns and do not have a predicate function.

Given that (34c)-type deverbal compounds are identified as VNs, their apparently complicated case patterns will follow from the case properties of their source verbs. Let us consider the deverbal compounds whose source verbs take two internal arguments, a theme and a locative. There are two options, depending on whether the left-hand element is interpreted as a theme argument or a locative argument, as shown in (39) and (40).

(39) Locative + transitive verb
     a. *X o [bin-zume]-suru* (X ACC [bottle-stuff]-do) 'to stuff X in a bottle'

     b. *X o [tana-age]-suru* (X ACC [shelf-put.on]-do) 'to shelve X' (metaphorical)

     c. *X o [huna-zumi]-suru* (X ACC [ship-load]-do) 'to load X on a ship'

(40) Theme + transitive verb

    a. *X ni [azi-tuke]-suru* (X LOC [seasoning-add]-do) 'to add seasoning to X'

    b. *X ni [dame-dasi]-suru* (X LOC [no.good-give]-do) 'to urge someone to do better'

    c. *X ni [ni-zumi]-suru* (X LOC [freight-load]-do) 'to load freight on X'

In (39), the source verbs, *tume-ru* 'stuff, put in', *age-ru* 'put on', and *tum-u* 'load, put on' are three-place predicates taking a theme and a locative argument in addition to the subject i.e. an agent. In these cases the locative arguments are compounded with the deverbal nouns, so that the theme arguments are realized in the accusative case externally. The examples in (40) display the reverse pattern, realizing the theme arguments internally and the locative arguments externally. Note that one and the same converted noun may exhibit two different case patterns, as shown by (39c) and (40c) involving *zumi* (< *tumu* 'load'). This fact confirms our assumption that deverbal compound formation of this kind is not constrained by a structural constraint like the ones proposed by Selkirk and Lieber (cf. Y. Yumoto 2014 for related discussion).

Another intriguing property of many deverbal compounds that function as VNs is what appears to be a double realization of the argument. Observe the examples in (41).

(41) a. *Gasorin   o   [ne-age]   si-ta.*
      gasoline   ACC   price-raise   do-PST
      'They raised the price of gasoline.'

    b. *Gasorin   ga   [ne-agari]   si-ta.*
      gasoline   NOM   price-go.up   do-PST
      'The price of gasoline went up.'

In (41a), the deverbal noun originating from a transitive verb *age-ru* 'to raise' is compounded with what appears to be its theme object *ne* 'price'. Nonetheless, the whole compound *ne-age* 'to raise the price' takes another object 'gasoline' which is realized syntactically as an accusative object. Since the verb 'raise' can take only one accusative object, a question arises as to where the "extra" object 'gasoline' comes from. Similarly, the intransitive compound *ne-agari* [price-rise] in (41b), consisting of the intransitive *agar-u* 'rise' and its subject *ne* 'price', takes the syntactic subject 'gasoline' marked with the nominative case. For convenience' sake, let us refer to the left-hand constituent of such a compound (e.g. *ne* in *ne-age*) as an "inside argument" and the noun that is manifested syntactically (e.g. *gasorin* 'gasoline') as an "outside argument".

In the literature two analyses have been proposed to account for this kind of relation between inside and outside arguments. On the one hand, Y. Yumoto (2010, 2014) points out that the inside and outside arguments in the peculiar argument realization pattern of (41) are characterized by a particular semantic relation in which the inside argument represents a "property" of the outside argument and determines its nature by specifying the price (*ne*), color (*iro*), width (*haba*), rank (*kaku*), and the like. Since the inside arguments require their possessor or holder as an argument, the compound inherits the "possessor/holder" role of the inside argument and realizes it as the outside argument on the clausal level. Compare the examples in (41) with those in (42) below. In (42a) and (42b), the genitive phrases *senzo no* 'of the ancestor' and *otoosan no* 'father's' appear to have the same kind of "possessor/holder" relation with the left-hand nouns (*haka* 'grave' and *kata* 'shoulder') within the compounds (Kageyama 1993; Chapter 14 [Kageyama, this volume]). Thus the construction of (42a, b) at first glance looks similar to that of (41).

(42) a. *senzo   no   [haka-mairi]*
          ancestor   GEN   [grave-visit]
          'a visit to one's ancestors' grave'

   a´. **senzo   o   [haka-mairi]-suru*
          ancesttor   ACC   [grave-visit]-do

   b. *otoosan   no   [kata-tataki]*
        father   GEN   [shoulder-tap]
        'giving Father a shoulder massage'

   b´. **otoosan   o   [kata-tataki]-suru*
          father   ACC   [shoulder-tap]-do

Compounds such as those in (42), however, are not VNs but simple event nominals because, as shown by the ungrammaticality of (42a´) and (42b´), they cannot function as predicates by being directly adjoined to *suru*. Because of this, the external nouns *senzo* 'ancestor' and *otoosan* 'father' can only be marked with the genitive *no* within the noun phrases headed by the compound nouns. Based on these observations Y. Yumoto (2010) develops a special mechanism of complex predicate formation for the (41)-type compound VNs using the Qualia Structures of a source verb and its inside argument.

The other analysis is suggested by Kageyama (1999) that makes use of the phenomenon called "doubling", where an identical or similar noun shows up at both word level and clause level, as in English denominal verb constructions like *He **shelved** the books on the top **shelf***. He notes that examples like those in (41) permit a noun identical to or synonymous with an inside argument to occur in duplicate externally, as far as redundancy does not result. Thus, the examples in

(43a, b) below, where the nominal suffix *-dai* meaning 'price' is added alongside the inside argument *ne* 'price', are as acceptable as those in (41a, b) that do not involve the suffix.

(43) a *Gasorin-**dai*** *o* *[**ne**-age]* *si-ta.*
     gasoline-price ACC price-raise do-PST
     'They raised the price of gasoline.'

    b *Gasorin-**dai*** *ga* *[**ne**-agari]* *si-ta.*
     gasoline-price NOM price-go.up do-PST
     'The price of gasoline went up.'

As pointed out in Kageyama (1980) and Chapter 3 (Kobayashi, Yamashita and Kageyama, this volume), a phenomenon analogous to (43) can be observed with Sino-Japanese VNs that consist of two morphemes. In (44), for example, where *to-zan* [climb-mountain] is a Sino-Japanese VN, the inside argument *zan* 'mountain' and the outside argument *Huzi-san* 'Mt. Fuji' have the same kind of semantic relation as do *ne* 'price' and *gasorin-dai* 'gasoline price' in (43).

(44) *Huzi-san* *ni* *[to-zan]-suru*
    Fuji-mountain DAT [climb-mountain]-do
    'to climb Mt. Fuji'

Note that (44) overtly exhibits the doubling of the morpheme 'mountain' (*san* and *zan*) on the inside argument and the outside argument. According to Taro Kageyama (p.c.), two-morpheme Sino-Japanese VNs appear to permit this kind of argument doubling more freely than native compounds like *ne-age*, but the underlying principle will be one and the same.

    Neither Kageyama (1999) nor Y. Yumoto (2010) offers a satisfactory account of exactly how the doubling phenomenon takes place and how the case marking on the "reduplicated" outside argument is available. If the inside argument *ne* 'price' in *ne-age* 'price raise' corresponds to the object of the source verb *ageru* 'raise', why can the whole compound additionally have the outside argument 'gasoline' marked with the accusative? The argument of the deverbal compound which functions as a VN is not necessarily marked with the accusative; dative marking is also observed as we saw in (44). This is an interesting question that cuts across the realms of morphology and syntax.

6 Conclusion and future research perspectives

This chapter has surveyed the complex properties of V-to-N and A-to-N conversion, with particular focus on the following two issues concerning deverbal and deadjectival compound words: (i) the mechanisms of semantic interpretation to explain

the diversity of meanings involved with conversion and deverbal compounding, and (ii) the conditions on the formation of complex event nominals by deverbal compounding. Concerning the first issue, we have argued that the diverse meanings of simplex and compound nouns involving conversion can be properly captured by distinguishing two types of nouns, entity nouns and event nouns. The distinction applies equally to both simplex and compound deverbal nouns. Except for special names due to creative word formation, the semantic interpretations of deverbal nouns can be accounted for by resorting to the lexical semantic representations of their source verbs, such as argument structure, LCS, and Qualia Structure. Concerning the second issue, we have observed that Japanese complex event nominals have intriguing properties that are not found with English deverbal compounds. Specifically, Japanese does not obey the structural constraint stipulated by Selkirk (1982) and Lieber (1983) that the internal argument of a verb must be preferentially satisfied inside the compound headed by the verb. Thus, *bin-zume* [bottle-stuff] can realize the object externally as in *wain o bin-zume-suru* [wine ACC bottle-stuff-do] 'to bottle the wine'. Furthermore, formation of a compound that can take an argument not selected by the source verb but provided from the left-hand noun (such as *ne-age* [price-raise] in *gasorin o ne-age-suru* [gasoline ACC price-raise-do] 'to raise the price of gasoline') is productive in Japanese. The difference between English and Japanese in these respects can be attributed to the fact that English deverbal compounds are nouns whereas Japanese deverbal compounds like *bin-zume* and *ne-age* are predicates (VNs) that can realize arguments syntactically.

Like all the other kinds of word formation and morphological processes, conversion and the compounding based on conversion call for elaborate systems of semantic interpretations including information on pragmatic knowledge. This chapter has briefly illustrated how the possible meanings of simplex and compound deverbal nouns can be formalized in terms of Lexical Conceptual Structure and Qualia Structure. In analyzing the meaning of compounds, the lexical meanings of not only the head verb but also the compounded element contribute to determining the semantic and syntactic properties of the whole compounds. It remains to be seen how the interaction of lexical meaning and word structure can be captured in a coherent theory and how the semantic properties of not only the source verb of the deverbal noun but also the left-hand element compounded with it are reflected in the syntactic and semantic properties of the compounds.

Acknowledgments

References

Allen, Margaret. 1978. *Morphological investigation*. Storrs, CT: University of Connecticut dissertation.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the definition of words*. Cambridge, MA: MIT Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Ito, Takane and Yoko Sugioka. 2002. *Go no shikumi to gokeisei* [Mechanism of words and word formation]. Tokyo: Kenkyusha.

Johnston, Michael and Federica Busa. 1996. Qualia structure and the compositional interpretation of compounds. In Evelyne Viegas (ed.), *Breadth and depth of semantic lexicons*, 167–187. Dordrecht: Kluwer.

Kageyama, Taro. 1980. *Goi no kōzō* [Lexical structures]. Tokyo: Shōhakusha.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation].Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1999. *Keitairon to imi* [Morphology and semantics] Tokyo: Kurosio Publishers.

Kageyama, Taro. 2001. Polymorphism and boundedness in event/entity nominalization. *Journal of Japanese Linguistics* 17. 29–57.

Kageyama, Taro. 2002. *Kejime no nai nihongo* [Japanese as an unbounded language]. Tokyo: Iwanami Shoten.

Kageyama, Taro. 2009. *Isolate*: Japanese. In Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Vocabulary strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kobayashi, Hideki, Kiyo Yamashita and Taro Kageyama. this volume. Chapter 3: Sino-Japanese words. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Lieber, Rochelle. 1983. Argument linking and compounding in English. *Linguistic Inquiry* 14. 251–286.

Makino, Seiichi. 1976. Nominal compounds. In Masayoshi Shibatani (ed.), *Syntax and semantics 5: Japanese generative grammar*, 483–498. New York: Academic Press.

Miyamoto, Tadao and Hideki Kishimoto. this volume. Chapter 12: Light verb constructions with verbal nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Namiki, Takayasu and Taro Kageyama. this volume. Chapter 6: Word structure and headedness. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Nishio, Toraya. 1988. *Gendai goi no kenkyū*. [A study of Japanese lexicon]. Tokyo: Meiji Shoin.

Okutsu, Keiichiro. 1975. Fukugō meishi no seisei bunpō [Generative grammar of compound nouns]. *Kokugogaku* 101. 19–34. Reprinted in Michiaki Saito and Masahiko Ishii (eds.), *Go kōsei* [Word composition], 158–175. Tokyo: Hituzi Syobo.

Ono, Naoyuki. this volume. Chapter 17: Agent nominals. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Ōno, Susumu. 1953. Nihongo no dōshi no katsuyōkei no kigen ni tsuite [On the origin of the inflection of verbs in Japanese]. *Kokugo to Kokubungaku* 30(6). 47–56.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pustejovsky, James, Pierrette Bouillon, Hitoshi Isahara, Kyoto Kanzaki and Chungmin Lee (eds.), 2013. *Advances in Generative Lexicon theory*. Dordrecht: Springer.

Selkirk, Elisabeth O. 1982. *The syntax of words*. Cambridge, MA: MIT Press.

Sugioka, Yoko. 1989. Haseigo ni okeru dōshi-sosei no uketsugi [Inheritance of verbal features in derived words]. In Susumu Kuno and Masayoshi Shibatani (eds.), *Nihongogaku no shintenkai* [New developments of Japanese linguistics], 167–185. Tokyo: Kurosio Publishers.

Sugioka, Yoko. 2001. Event structure and adjuncts in Japanese deverbal compounds. *Journal of Japanese Linguistics* 17. 83–108.

Sugioka, Yoko. 2002. Incorporation vs. modification in Japanese deverbal compounds. In Noriko Akatsuka and Susan Strauss (eds.), *Japanese/Korean Linguistics* 10. 496–509. Stanford: CSLI Publications.

Williams, Edwin. 1981. On the notions 'lexically-related' and 'head of a word'. *Linguistic Inquiry* 12. 245–274.

Yumoto, Shōnan. 1978. *Awase meishi no imi kijutsu o megutte* [On semantic interpretation of compound nouns]. In Hirotake Matsumoto (ed.), *Nihongo kenkyū no hōhō* [Methods of Japanese language studies], 75–93. Tokyo: Mugi Shobō.

Yumoto, Yoko. 2010. Variation in N-V compound verbs in Japanese. *Lingua* 120. 2388–2404.

Yumoto, Yoko. 2014. "Meishi + dōshi" gata fukugōgo ga jutsugomeishi to naru jōken [Conditions for predicative use of N-V compound nouns]. In Hideki Kishimoto and Yoko Yumoto (eds.), *Fukuzatsu jutsugo kenkyū no genzai* [Current issues in complex predicate research], 179–203. Tokyo: Hituzi Syobo.

Yumoto , Yoko. 2015. "Meishi + dōshi" fukugōgo no tōgohanchū to imiteki kategorī [Syntactic category and semantic category of N-V compound nouns.] In Takashi Masuoka (ed.) *Nihongokenkyū to sono kanōsei* [Japanese-language studies and their potentiality], 80–105. Tokyo: Kaitakusha.

Yoko Sugioka and Takane Ito

10 Derivational affixation in the lexicon and syntax

1 Introduction

This chapter deals with word derivation by affixation, which, along with compounding, is one of the most active and productive word formation processes. Japanese employs two types of affixation, prefixation as in (1a) and suffixation as in (1b).[1] No clear case of infixation is known.

(1) a. prefixation: *ko-isi* 'pebble' < *isi* 'stone'

    b. suffixation: *yowa-sa* 'weakness' < *yowa(i)* 'weak'

In the field of traditional Japanese grammar, where morphology and syntax (grammar) are held to be separate areas of inquiry, research on word formation occupies a rather marginal place compared with grammatical research. Among the works available, Sakakura (1966) is considered a seminal work providing a detailed description of derivational affixes in the native Japanese stratum on the basis of examples taken mainly from Old Japanese, where there was little or no influence of borrowings from Chinese or any other language. Saito and Ishii (eds.) (1997) survey the literature in this area, tracing the history of word formation studies in the domestic tradition. The theoretical approaches to Japanese morphology inspired by more recent work on English and other European languages, on the other hand, brought to light a number of interesting phenomena that would pose problems for the general theory of morphology because of the agglutinative character of the language. Because issues related to compounding are discussed in other chapters of this volume, the present chapter focuses on affixation and clarifies in what way Japanese may potentially contribute to the theorization of word formation phenomena.

Derivational affixation by nature involves multiple components of grammar: phonology with respect to changes in form, morphology with respect to lexical categories and productivity, semantics with respect to the compositional and holistic interpretations of the output, and syntax with regard to the function in syntactic structure and argument realization. As a matter of fact, the place of morphology in grammar has been a central topic of heated controversy in generative grammar since

---

1 In presenting examples, morpheme by morpheme glosses are omitted when translation provides enough information as in (1). So called citation forms instead of verbal/adjectival stems are given, with either *ru/ u* 'PRS' for verbs or *i* 'PRS' for adjectives added in parentheses.

Chomsky (1970), and the long-standing theoretical debate between strong lexicalism and strong syntacticism still continues. Various theoretical models that have been proposed for the role and architecture of derivational processes in particular and morphology in general can be classified roughly into three competing approaches. First, the syntactic approach attributes all productive word formation processes to syntax, as proposed in Lieber (1992) and in more recent frameworks of Distributed Morphology (Halle and Marantz 1993) and L-syntax approach (Hale and Keyser 1993), among others. Second, the lexical approach relegates all word formation processes to the lexicon, as in Di Sciullo and Williams (1987) and in non-derivational frameworks like Head-Driven Phrase Structure Grammar (HPSG: Pollard and Sag 1994).[2] The third is a hybrid or modular approach where word formation is supposed to take place in both syntax and the lexicon under different conditions (Kageyama 1989; Shibatani and Kageyama 1988; Jackendoff 1997, 2002; Sadock 2012).

One of the prominent characteristics of Japanese as an agglutinative language is that clause-taking predicates are realized as suffixes in syntactic structure. For example, the predicate *hatarak-ase-rare-ta-gar(u)* [work-CAUS-PASS-DESI-express] 'express eagerness to be forced to work' consists of five morphemes (excluding the non-past or present tense inflection *-u* at the end) that are hierarchically organized by multiple embeddings in syntactic structure. It is no wonder that the syntactic approach to morphology has been widely practiced by those who study these syntactic phenomena ever since the rise of generative studies of Japanese (Kuroda 1965; Kuno 1973; Shibatani 1976), and the intricate nature of the interaction of syntax and morphology in Japanese has been instrumental in the development of a hybrid or modular approach (Sugioka 1986; Kageyama 1993), in the study of affixation as well as compounding (see Chapter 8 [Kageyama, this volume] for the distinction of lexical vs. syntactic V-V compounds).

It has long been noticed that affixes can be categorized into different classes, depending on their phonological/morphological regularity and productivity, and that more regular/productive affixes are placed outside less regular/productive ones, as explicitly stated in Siegel's (1974) Level Ordering Hypothesis. The issue has been couched in X-bar type theory, where the more regular/productive affixes are attached to a larger unit (say, stem or word) while the less regular/productive ones are to a smaller unit (say, root) (Selkirk 1982). Kageyama (1993, 2001) argues, for instance, that Japanese has four different classes of building blocks for word-formation: root, stem, word, and word⁺. A root is defined as an unanalyzable mono-morphemic unit, while a stem is a mono-morphemic or poly-morphemic base to which an inflectional affix can be attached. Thus in an adjective *huka(i)* 'deep', *huka* is a root and at the same time a stem, while in an adjectival compound *oku-huka(i)* [end-deep] 'profound' *huka* is a root, and *oku-huka* is a stem. The distinction between root and stem is necessary for capturing the difference in regularity/productivity of affixes

---

2 Lexical approaches were also proposed in Japanese, as in Farmer (1980) and Miyagawa (1989).

that attach to a unit smaller than a word, like the non-productive nominalizing suffix *-mi* and the productive *-sa*, which are attached at the root level and the stem level, respectively, but not at the word level (*huka-mi/huka-sa*; **oku-huka-mi/oku-huka-sa*; note that the base is not a free-standing word: see Section 4.2 for details). Furthermore, Kageyama (2001) argues for the category "word+", which is larger than a word but is not a syntactic phrase. Affixation to a word+ exhibits phrase-like properties (e.g., phrasal accent pattern, incorporation of conjoined elements) while obeying lexical integrity as a whole: the prefix *zen-* 'ex-' for instance selects for a word+; it requires a slight pause after it (designated by "|"), has its own accent with the accent pattern of the base intact (*ZEn|soORI-DAizin* 'ex-Prime Minister'), and can be attached to a conjoined sequence of words (*zen [soori-daizin oyobi gaimu-daizin]* 'ex-Prime Minister and ex-Foreign Minister'); the prefix, however, cannot be attached to a syntactic phrase (**zen [kaisya no syatyoo]* (ex [company GEN president]) 'ex-company's president').

Furthermore, some regular and productive word formation processes can apply to phrases, and in some cases even involve sentence-level embedding, with many "clause-taking" predicates realized as suffixes (Kageyama 1993). For instance, aspectual suffixes like those exemplified in (2) are attached to VPs (see Section 4, Table 4 (k-n)). It should be noted that the case-marking pattern can be either that of VPs as in (2a) or that of NPs as in (2b) with the genitive marker *no* (see Section 4.1):

(2) a. *[Hokkaidoo o ryokoo]-tyuu* b. *[Hokkaidoo no ryokoo]-tyuu*
        Hokkaido ACC travel-while   Hokkaido GEN travel-while
        'while travelling in Hokkaido'

Also, some predicate-forming suffixes like *-sase* take clauses as their complements, as will be discussed in Section 5.2. While these suffixes take phrases/clauses as their basic usage, some noun-taking affixes can take a phrase-like element in extended usage. For example, the Sino-Japanese suffix *-huu* 'style' selects for a simple noun in its basic use as in *wakamono-huu* [youngster-style] 'a style typical of youngsters', but can exceptionally take an NP as its scope as in *[tokai no wakamono]-huu* ([city GEN youngster]-style) 'a style typical of urban youngsters'.

Related to the question of what kinds of morphological units are involved in derivation (root, stem, word, word+, phrase) is the perennial issue of the place of morphology in the theory of grammar, already mentioned briefly above. Can words be derived in syntax in the same way that phrases are derived (as in the theory of Distributed Morphology), or are they all produced and stored in the lexicon, with generalizations on morpheme combinations captured in the lexical component in a different way from syntactic derivation (as in strong lexicalism)? Since phrases are by definition generated in syntax, the existence of affixes that are attached to phrases (e.g. aspectual affixes shown in (2), and predicate-forming suffixes to be discussed in Sections 3 and 5) provides support to the syntactic analysis of some types of word derivation. On the other hand, affixes that select smaller units tend to

have low productivity with many unpredictable lexical gaps, suggesting the plausibility of lexical treatment of such types of word derivation. Hence the observation of various affixation processes offers supporting evidence for a hybrid or modular approach to morphology.

This chapter surveys various affixation processes in Japanese with special focus on the contrast between their lexical and syntactic properties. Because of its agglutinative nature, Japanese offers various ideal cases for delving into the dual nature of word formation. Table 1 shows major diagnostics of the syntactic and lexical processes of affixation:

Table 1: Characteristics of syntactic/lexical affixation

	syn. affixes	lex. affixes
(a) can attach to phrase	YES	NO
(b) can attach to syntactically derived word (syntactic compound/word with syntactic affix)	YES	NO
(c) lexical gaps	NO	YES
(d) non-compositional/lexicalized meaning	NO	YES

We take (a) and (b) to be the most compelling evidence for the syntactic status of an affix, while (c) and (d) provide strong support for the lexical status.[3] It should be noted that the terms "syntactic" and "lexical" in this chapter refer to the components of grammar to which the relevant affixation is ascribed rather than the level of the base (namely phrase or word), which distinction we believe to bear direct relation to the psychological status of being generated on-line or being memorized in the lexicon. Hence the characteristics in Table 1 constitute sufficient, but not necessary, conditions of the relevant status: for instance, an affix that attaches to a word (not a phrase) can be syntactic if it exhibits full productivity and regularity and can attach to syntactically derived words.

The issue of syntax-lexicon dichotomy has also been a topic of hot debate in psycho- and neurolinguistic studies, often dubbed the "past tense debate" (Pinker and Ullman 2002; among many others), where it has been discussed whether or not two different mental mechanisms (associative memory and computation by symbolic

---

3 Note, however, that words with a syntactic affix can be lexicalized. For instance, the noun *waru sa* 'mischief', consisting of an adjective *waru(i)* 'bad' and the suffix *sa*, has an unpredictable "lexicalized" meaning. This does not, however, undermine the claim that *sa* is a syntactic affix as will be argued in Section 4.2 based on the above diagnostics. High frequency words containing a syntactic affix may be memorized, and may undergo semantic change to exhibit non compositional semantics.

rule) supported by different neural mechanisms are employed in word derivation/inflection.

In the following four sections, we will examine various word formation processes: prefixation (Section 2), adjective formation (Section 3), nominalization (Section 4), and verb formation (Section 5). Section 6 will present some supporting evidence for the dual nature of word formation processes from neurolinguistic experiments. The last section provides concluding remarks.

A brief note on the coverage of this chapter. Affixes may be *wago* (native Japanese) or Sino-Japanese (S-J). While they mostly attach to the bases of the same type, some *wago* affixes (e.g. *-sa*) can attach to S-J bases, and some S-J affixes (e.g. *tyoo-*) can attach to *wago* bases. Although derivation by S-J morphemes are not discussed in detail in this chapter, some are taken up because of this mixing of lexical strata in derivational affixation (see Chapter 3 [Kobayashi, Yamashita and Kageyama, this volume] on S-J morphemes). Note also that when category-changing affixes in Japanese attach to inflecting words (e.g. verbs, adjectives), they attach to "stems" rather than to full-fledged, inflected words. Assuming this general property of affixes, we will simply say the suffix *-sa*, for example, attaches to "adjectives", when strictly speaking, we should say "adjectival stems" or "stems of adjectives".

2 Prefixation

This section surveys the basic properties of major prefixes (as exemplified in Table 2) that attach to *wago* bases.

Prefixes tend to be non-category determining: one and the same prefix may attach to various categories, deriving words of the same category as the base words. For instance, a diminutive prefix *ko-* 'little' attaches to an N, A, AN, and V:

(3)   a.   N > N: *tori* 'bird' > *ko-tori* 'small bird'

       b.   A > A: *taka(i)* 'high' > *ko-daka(i)*[4] 'slightly elevated'

       c.   AN > AN: *iki (na)* 'refined' > *ko-iki* (na) 'stylish'

       d.   V > V: *tuk(u)* 'push' > *ko-zuk(u)* 'poke'

---

4 The initial consonant of the base /t/ is changed into /d/ by a well known phonological process called *rendaku* (sequential voicing), which applies to the initial obstruent of the second member of a *wago* complex word (Vance 1987). *Rendaku* is most productively observed in compound nouns, but is also found in some prefixed words.

Table 2: Prefixes in Japanese

prefix		meaning	base cat.	output cat.	examples	comments
Negative prefixes						
(a)	*hu* (S J)	negation	AN/N/ VN/root	AN/N	*hu tasika (na)* 'uncertain'; *hu ninki (na)* 'unpopular'; *hu bi* 'inadequacy'	semi productive
(b)	*hi* (S J)	negation	AN/N/ root	AN/N	*hi ningen teki* 'inhuman'; *hi un* 'misfortune; *hi bon (na)* 'unusual'	semi productive
(c)	*mi* (S J)	incomplete	VN/root	AN/N	*mi happyo (no)* 'unpublished'; *mi kan (no)* 'unfinished'	semi productive
(d)	*mu* (S J)	negation, absence	N/root	AN/N	*mu imi (na)* 'meaningless'; *mu kyuu (no)* 'unpaid'	semi productive
(e)	*bu* (S J)	negation, absence	N/root	AN	*bu situke (na)* 'ill mannered'; *bu sui (na)* 'unrefined'	semi productive
Degree prefixes						
(f)	*tyoo* (S J)	super	N/A/AN	same as the base	*tyoo tensai* 'super genius'; *tyoo kawai(i)* 'super cute; *tyoo mazime (na)* 'super serious'	fully productive in colloquial style
(g)	*oo*	large (in quantity/ degree)	N/AN/VN	same as the base	*oo yuki* 'heavy snow'; *oo mazime (na)* 'very serious'; *oo abare (suru)* 'rampage'	semi productive
(h)	*ko*	diminutive, sometimes pejorative	A/N/V/ AN/VN	same as the base	*ko zakasi(i)* 'shrewd'; *ko yakunin* 'petty official'; *ko zuk(u)* 'poke'; *ko girei (na)* 'clean'; *ko odori (suru)* 'jump for joy'	semi productive
(i)	*do*	emphatic, often pejorative	N/A/ AN/V	same as the base	*do sirooto* 'utter layman'; *do gitu(i)* 'shocking'; *do hade (na)* 'gaudy'; *do tuk(u)* 'punch'	unproductive, used only in colloquial style
(j)	*han* (S J)	half, semi	N/VN/AN	same as the base	*han tosi* 'half a year'; *han naki (no/suru)* 'half crying; *han toomei (na/no)* 'semi transparent'	unproductive
(k)	*nama*	slight, insufficient	N/A/ AN/VN	same as the base	*nama akubi* 'slight yawn'; *nama atataka(i/na)* 'lukewarm'; *nama kaziri (no/suru)* 'smattering'	unproductive, cf. *nama (no)* 'raw' (*nama zakana* 'raw fish')

Table 2: (continued)

prefix		meaning	base cat.	output cat.	examples	comments
Other prefixes						
(l)	*o*	honorific, beauific	A/N/AN/ VN/V	same as the base	*o utukusi(i)* 'beautiful'; *o kane* 'money; *o kirei (na)* 'beautiful'; *o torihiki (suru)* 'make a deal'; *o wasure ni.nar(u))* 'forget'	productive with *wago*; fully productive in the honorific construction
(m)	*ma*	pure, exact	N/A/AN	same as the base	*ma mizu* 'pure water'; *ma atarasi(i)* 'brand new'; *ma syooziki (na)* 'utterly honest'	unproductive
(n)	*ki*	pure	N/AN	same as the base	*ki zyooyu* 'pure soy sauce'; *ki mazime (na)* 'very serious'	unproductive
(o)	*su*	bare	N/VN/ root	same as the base	*su asi* 'bare foot'; *su doori suru* 'pass through without stopping'; *su nao (na)* 'docile'	unproductive
(p)	*zen* (S J)	the last, previous	N	N	*zen sai* 'ex wife'; *zen syusyoo* 'the last Prime Minister	productive with nouns denoting official posi tions, phrasal accent
(q)	*moto*	ex , former	N	N	*moto tuma* 'ex wife'; *moto otto* 'ex husband'; *moto syusyoo* 'ex Prime Minister'	productive with nouns denoting official posi tions, phrasal accent

This is in sharp contrast with suffixes, which tend to determine the category of derived words. For example, *-r* derives a verb from a noun or an adjective as in (4a,b), while *-sa* derives a noun from an adjective or an adjectival noun as in (4c,d):

(4) a.   N > V: *ziko* 'accident' > *ziko-r(u)* 'meet an accident'

  b.   A > V: *huto(i)* 'thick' > *huto-r(u)* 'become fat'

  c.   A > N: *yasasi(i)* 'gentle' > *yasasi-sa* 'gentleness'

  d.   AN > N: *sinsetu (na)* 'kind' > *sinsetu-sa* 'kindness'

Since the head of a word determines the category, this contrast suggests that the head of a word is, in principle, on the right-hand side of a word. The contrast is observed not only in Japanese, a typical head-final language, but is shared by many other languages, a generalization known as "Right-hand Head Rule" (Williams 1981: see Chapter 6 [Namiki and Kageyama, this volume]).

As will be discussed in Section 4.1, V-to-N conversion is not productive in simple words, so that many verbs in their *ren'yō* (infinitive) forms cannot be used either as a noun or as a VN (*kaziri no [chew GEN] / *kaziri suru [chew do] < *kazir(u)* 'chew'), but prefixation, as well as compounding (Chapter 9 [Yumoto, this volume]), tends to rescue, as it were, conversion of verb infinitives to nouns or VNs, as in *nama-kaziri no* [half-chew GEN] 'half-understood', *nama-kaziri suru* 'smatter' (see Table 2(g,h,j,k)).

Although most prefixes attaching to smaller units have low productivity, a few prefixes with a phrasal accent pattern are quite productive in a specified semantic domain: *zen-* and *moto-* (Table 2(p,q)), as discussed in Section 1, select for a word⁺ category, and are productively attached to nouns denoting official positions. Also, *tyoo-* (Table 2(f)), which is fully productive in colloquial language, exhibits a phrasal accent pattern (e.g., *TYOo | iSOGASI(i)* 'extremely busy'), even though the same prefix shows a word-level accentuation in non-colloquial S-J affixation (e.g. *tyoO-TOKkyuu* 'super-express train').

Another prefix that shows high productivity is *o-* (Table 2(l)), which can be attached to native Japanese words of any category to derive a beautific/honorified form, though some pragmatic factors seem to be involved in judging its acceptability. For example, it is pragmatically unlikely for the beautific prefix to be added to nouns like *hokori* 'dust' (*o-hokori). This prefix exhibits full productivity with verb infinitives in the so-called honorific construction, as in (5), where the referent of the subject is the target of honorification. The prefix *o-* is interesting in that it can attach not only to lexical words but also to verbs containing syntactic affixes like causative *-sase* (see Section 5.2), as shown in (5b), a fact that lends credence to the syntactic status of this prefix.

(5) a. *Sensei   ga   hon   o   o-motome-ni.nar-u.*
       teacher   NOM   book   ACC HON-buy-HON-PRS
       'The teacher buys a book.'

   b. *Sensei   ga   gakusei   ni   hon   o   o-yom-ase-ni.nar-u.*
       teacher   NOM   student   DAT   book   ACC HON-read-CAUS-HON-PRS
       'The teacher makes students read the book.'

There is a phonological constraint prohibiting the honorific prefixation from applying to a mono-moraic base, as shown by the ungrammaticality of **o-ni-ni.nar(u)* < *ni(ru)* 'boil' or 'resemble', **o-mi-ni.naru* < *mi(ru)* 'take a look', and **o-ki-ni.naru* < *ki(ru)* 'wear' or *ku(ru)* 'come'.[5]

---

5 It might be the case that the last two examples are blocked by the dedicated honorific forms (*goran ni.nar(u)* 'take a look' and *o mesi ni.naru* 'wear' or *irassyar(u)* 'come'). The blocking effect, however, is not perfect: *o iki ni.nar(u)* 'go' is sometimes used in the colloquial style in spite of the existence of its honorific form *irassyar(u)*. Thus the mono moraic constraint is significant in ruling out these examples.

In sum, prefixes in Japanese exhibit varying degrees of productivity, with a high correlation between productivity and the level of affixation: Those prefixes that are attached to larger units of word+ and words with syntactic affixes are distinctly more productive than those that attach to smaller units.

Note in passing that an interesting mismatch between morphological structure and semantic interpretation can be caused by prefixation. A notable example is the prefix *ko-* 'little' (Table 2(h)) as it is used in the idiomatic phrase of (6). Morphologically, it is attached to the following noun *kubi* 'neck', whereas in semantic interpretation, as suggested by the English translation, the prefix takes the VP (*kubi o kasige(ru)* 'bend one's head') as its scope. The expression in (6) thus does not refer to the size of one's neck, but rather the degree to which the head is bent.

(6) *ko-kubi   o   kasige-ru*
     little-neck   ACC   bend-PRS
     'to bend one's head slightly to one side (as if questioning the truth)'

Kitagawa (1986) proposes to account for this kind of mismatch by assuming that the prefix, originally attached to the base noun, syntactically moves up to a higher structure to take scope over the VP in LF. On the other hand, Fukushima (2014) attempts to give a formal semantic analysis of the idiomatic expressions at issue from the lexicalist perspective. No matter whether one takes a syntactic approach or a semantic approach, however, it is important to note that *ko-* prefixation with this kind of mismatch has absolutely no productivity, allowing only a very small number of fixed expressions (compare (6) with the unacceptable **ko-kubi o nobas(u)* [little-neck ACC stretch] 'to stretch one's neck a little') often accompanied with highly idiomatic meanings, as in *ko-mimi ni hasam(u)* [little-ear DAT put.in] 'overhear'). Thus this phenomenon does not fit the general characteristics of syntactic/lexical affixes discussed in Section 1 in that its low productivity points to its lexical status despite its seemingly phrasal interpretation.

3 Adjective-forming suffixes

As shown in Table 3, Japanese abounds with adjective- and AN-forming suffixes in both the *wago* and Sino-Japanese strata. This section surveys their characteristics with particular reference to the lexical-syntactic distinction.

Table 3: A/AN-forming suffixes in Japanese

	suffix	base cat.	meaning	examples	comments
Affixes denoting subjectivity					
(a)	*asi(i)*	V	causing the emotion of V	*urayam asi(i)* 'enviable'; *yorokob asi(i)* 'delightful'	semi productive, selects emotive verbs
(b)	*si(i)*	redupli cated A, N	showing the property of A, N	*yowa yowa si(i)* 'weak looking'; *mizu mizu si(i)* 'juicy'	semi productive
Affixes denoting appearance/property					
(c)	*ppo(i)*	N	showing property of N	*otoko ppo(i)* 'masculine looking'; *abura ppoi* 'oily'	semi productive, idio syncratic meaning
(d)	*rasi(i)*	N	possessing the charac teristics of N	*otoko rasi(i)* 'manly'	productive, homophonous with auxiliary 'seem'
(e)	*kusa(i)*	N, A, AN	showing undesirable property of N, A, AN	*inaka kusa(i)* 'rustic'; *huru kusa(i)* 'arcane'; *keti kusa(i)* 'stingy'	unproductive, cf. *kusa(i)* 'smell bad' (*koge kusai* 'smelling burned')
(f)	*ge (na)*	A, AN	showing property of A, AN	*uresi ge* 'happy looking'; *manzoku ge* 'satisfied looking'	semi productive, cf. N suffix *ge* 'appearance' (*otona ge* 'adult property')
(g)	*soo (na)*	A, AN, V	look like A, AN, seem to V	*oisi soo* 'delicious looking'; *genki soo* 'healthy looking'; *ari soo* 'seems to exist'	productive, mostly used as auxiliary
(h)	*teki (na)* (S J)	N	possessing the property of N	*ningen teki* 'humane'; *ura banasi teki* 'like an inside story'	productive with S J N; semi productive with native N, loan N, compound and NP in colloquial style
VP taking affixes					
(i)	*ta(i)*	V	wanting to V	*mi ta(i)* 'want to see'	fully productive
(j)	*yasu(i)*	V	easy to V	*mi yasu(i)* 'easy to see'	fully productive
(k)	*niku(i)/ zura(i)/ gata(i)*	V	hard to V	*mi niku(i)* 'hard to see'	fully productive

3.1 Lexical adjectival suffixes

First, the suffixes *-asi(i)* and *-si(i)*, dating back to Old Japanese, are found in many subjective adjectives (Sakakura 1966; Yamazaki 1992). The more transparent of the two, *-asi(i)* can attach to a verb root to derive an adjective denoting an emotion or a subjective evaluation and induces a case marking alternation by marking the

object NP of the base verb with the nominative case as the subject of the derived adjective, in a way similar to derivation of evaluative adjectives with *-able* in English (e.g. *desire > desirable*):

(7) a. *syoosin   o   nozom-u   / syoosin   ga   nozom-asi-i.*
      promotion ACC desire-PRS   promotion NOM desire-able-PRS
      'to desire promotion / Promotion is desirable.'

   b. *yorokob(u)* 'rejoice' / *yorokob-asi(i)* 'delightful'; *utagaw(u)* 'doubt' / *utagaw-asi(i)* 'suspicious'; *konom(u)* 'like' / *konom-asi(i)* 'likeable'

The other affix *-si(i)* can attach to reduplicated A or N roots (*waka$_A$-waka-si(i)* [young-young-looking] '(very) young-looking'; *doku$_N$-doku-si(i)* [poison-poison-looking] 'poisonous-looking') with an emphatic meaning. Neither *-asi(i)* nor *-si(i)* attaches to phrases; they are only semi-productive, so that the affixed adjectives as a whole are listed in the lexicon.

3.2 Adjectival suffixes denoting appearance

The second group of A-forming affixes, *-ppo(i)* and *-rasi(i)* denote a property or the appearance of what is denoted by the base word, but they differ in connotation. While *N-ppo(i)* denotes a superficial and often undesirable property, *N-rasi(i)* denotes a prototypical or desirable property, as exemplified by the contrast between *kodomo-ppo(i)* 'childish' and *kodomo-rasi(i)* 'childlike'. Only the former can be used to modify an entity of a different category (e.g. *kodomo-ppo(i) otona* 'childish adult' vs. **kodomo-rasi(i) otona* 'childlike adult'). The cognitive meanings and connotations of the adjectives derived by these suffixes show idiosyncrasies, and there is no apparent reason to consider their derivation as syntactic.

    On the other hand, the same two affixes, when functioning as auxiliaries that are attached to tensed clauses, as in (8), express the modality of prediction or inference by the speaker, which may be called "evidentials" (cf. Teramura 1984: 224).

(8) *Ame ga   hut-ta   {rasi-i   / ppo-i.}*[6]
    rain NOM fall-PST {seem-PRS / appear-PRS}
    'It looks like it has rained.'

---

6 The use of *ppo(i)* as an auxiliary attaching to a tensed form is found only in casual colloquial speech.

These examples lend support to a hybrid (modular) theory of word formation where one and the same form can have a dual function of a derivational affix in the lexicon and an auxiliary-like affix in syntax.[7]

There are AN-forming suffixes that also function as auxiliaries that mark evidentiality: *-ge* derives ANs that denote appearance or some sensation from the point of view of the speaker (e.g. *ayasi-ge* 'suspicious-looking'). ANs derived by *-soo* denote some properties inferred by the speaker (e.g. *muzukasi-soo* 'difficult-looking'). They differ in their selection of the base in that *-ge*, which is more archaic-sounding, attaches only to a subclass of adjectives while *-soo* attaches to verbs as well: e.g. *naki-soo* /*naki-ge* 'looks like (s/he is) going to cry'. Furthermore, *-soo*, but not *-ge*, is capable of attaching additionally to tensed verbs with the meaning of evidentiality of reporting: *kaet-ta soo (da)* [return-PST EVID (COP)] '(I) hear that (s/he) came back'. Nevertheless, they both serve an important function in Japanese, where evidential markers are mandatory when adjectives denoting inner feelings are predicated of 3rd-person subject NPs as shown by the contrast below.

(9) a. *Ken wa   kanasi-{ge/soo}   da.   / *Watasi wa   kanasi-{ge/soo}   da.*
    Ken TOP sad-{look/seem} COP / *I   TOP sad-{look/seem} COP
    'Ken {looks/seems} sad / *I {look/seem} sad.'

   b. *{Watasi / *Ken} wa   kanasi-i.*
     {I   / *Ken} TOP sad-PRS
     '{I am / *Ken is} sad.'

It is a prominent feature of Japanese that emotive adjectives as exemplified above have a person restriction on the subject NP (Kuroda 1973). The fact that both *-ge* and *-soo* affect the choice of the subject NP indicates that they function as auxiliaries in syntactic structure; this contention is consistent with the fact that both can attach to syntactically derived adjectival phrases (e.g. *VP-ta(i)*; see Section 3.3) while preserving the case marking in VP, as shown below.

(10) *[[monku   o   ii]$_{VP}$-ta]$_{AP}$-{ge   / -soo}   na   gakusei*
    complaint ACC say-DESI-{look / seem} COP student
    'a student who seems to want to say a complaint'

---

7 These affixes, as well as the AN forming suffix *soo* discussed below, can be viewed as clitics (enclitics) in their auxiliary like usage attaching to tensed verbs, although they (except for *ppo(i)*) maintain their independent accent patterns and thus do not form a prosodic unit with the preceding element like clitics in the traditional sense. See Anderson (1992: Chapter 8) for variations of clitics in their phonological and syntactic properties.

   It may also be possible to view the development of the use of *ppo(i)* and *rasi(i)* as clitics as an instance of grammaticalization, in which a derivational suffix in the lexicon acquired the status as a sentential clitic (i.e. auxiliary). If so, it would provide counterexamples to the standard theory of grammaticalization (Heine and Kuteva 2002; and others), according to which grammaticalization proceeds unilaterally from a word to a clitic to an affix (Taro Kageyama, p.c.).

We can assume that as auxiliaries -*ge* and -*soo* are positioned "outside" the AP in syntax and form a word (i.e. *ii-ta-ge*) in phonological structure, hence their attachment to the base do not affect the syntactic structure within AP.

Consequently, adjectival affixes denoting appearance in Japanese include lexical affixes on one hand, and the ones that function as auxiliaries and are syntactic on the other. Some affixes serve as both, which may make the delineation seem unclear. However, one way to distinguish derivational affixes from auxiliaries is the possibility of further derivational affixation. Thus, the derivational affix -*ppo(i)* and -*rasi(i)* attaching to N to denote some property can be nominalized by the suffix -*sa* without any problem (*kodomo-ppo-sa*, *kodomo-rasi-sa* 'childlike-ness'), but if the same affix -*rasi(i)* is used as an auxiliary it cannot be nominalized: **yuuzai rasi-sa* [guilty looking-ness] 'seeming to be guilty'. In a similar vein, ANs ending with -*soo* and -*ge* resist nominalization by -*sa* (**kanasi-soo-sa*, ??*kanasi-ge-sa* 'sad-looking-ness'), which conforms to the observation above that they function as auxiliaries rather than derivational affixes.

3.3 Adjectival suffixes attaching to VP

The last group of suffixes including -*ta(i)* 'want to', -*yasu(i)* 'be easy to', -*niku(i)/ -zura(i)/-gata(i)* 'be hard to' (Table 3(i,j,k)) is unequivocally syntactic in that they attach to VP, giving rise to desiderative and "tough" predicates (see Sugioka 1986; Kuroda 1987; and Inoue 1978; inter alia for the details of the syntactic constructions involving these affixes). Their syntactic status is corroborated by their ability to attach to causative and passive verbs, which are syntactic in and by themselves.

(11) *[Sensei ni   home-rare]-{ta-i   / -yasu-i}.*
    teacher DAT praise-PASS-{DESI-PRS / -easy-PRS}
    '{(I) want to be / (I am) easily} praised by the teacher.'

These syntactic affixes can induce a case-marking alternation on the object NP of the base verb, from accusative to nominative, in accordance with the case marking pattern of adjectives (i.e. simple adjectives cannot take accusative case): hence in such cases we can assume that the verb is attached to the affix by head-movement in the syntactic structure and as a result cannot assign the accusative case on the object NP, thereby inducing the nominative marking on the object NP as in (12b).

(12) a. *Pan   o   tabe-ru.*   b. *Pan   ga   [tabe-ta-i   / -yasu-i]$_A$.*
       bread ACC eat-PRS   bread NOM eat-DESI-PRS / easy-PRS
       '(I) eat bread.'   '{(I) want / (It) is easy} to eat bread.'

This head-movement of the verb is optional for *-ta(i)* (cf. *Pan o tabe-tai* [bread ACC eat-DESI]), while it is obligatory for the "tough" affixes that denote the ease or difficulty of action. When "tough" affixes express the likelihood of an event denoted by VP, however, there is no case-marking alternation, which indicates that the affixation leaves the syntactic structure within the VP intact, resulting in a "mixed" category; i.e. AP constituting of VP and an adjectival affix (see also (11)).

(13) *[Kodomo ga kega o si]$_{VP}$-yasu$_A$-i.*
     child   NOM injury ACC do-easy-PRS
     'It is likely that children get an injury.'

Although the adjectives they derive may form a structure that contains the case marking pattern of VP as in (13), it is not plausible to see them on a par with auxiliaries like *-soo* discussed in Section 3.2. This is because they can take additional derivational affixes, such as nominalizing *-sa* (*yomi-ta-sa* [read-DESI-NMNL] 'desire to read', *kega no si-yasu-sa* [injury GEN do-easy-NMNL] 'likelihood of getting an injury') and verbalizing *-gar(u)* (*yomi-ta-garu* 'act like wanting to read': see Section 5.2). This observation supports the contention that the "tough" and desiderative affixes are phrasal affixes attaching in syntax and should be considered derivational affixes as opposed to auxiliaries like *-soo*, which are presumably attached to the base in phonological structure rather than in syntax.

To recapitulate, while some adjectival affixes are semi-productive and lexical in nature, there are other A- or AN-forming affixes that can attach to phrases and so are syntactic in nature. Furthermore, some affixes that denote appearance or speaker's inference function as modal auxiliaries, and can in some cases attach to tensed clauses. The flexibility of these A- and AN-forming affixes in the level and scope they cover reveals the agglutinative nature of Japanese predicate formation.

4 Nominalizing suffixes

Table 4 shows a list of suffixes that change verbs and adjectives into nouns. Among these noun-forming suffixes, this section will focus on the properties of two suffixes, one deverbal and the other deadjectival, and illuminate their syntactic nature by contrasting them with other bona fide lexical affixes of nominalization.

Table 4: Noun-forming affixes in Japanese

	suffix	base cat.	meaning	examples	comments
Conversion nominals					
(a)	no suffix	V	act/result/ person	*aruki* 'walk'; *tutumi* 'bundle'; *toori* 'passage'; *yopparai* 'drunkard'	semi productive (productive with compound V)
Manner nominals					
(b)	*kata*	V	way of V ing	*kaki kata* 'manner of writing'; *ari kata* 'way of being'; *hanasi hazime kata* 'how to start speaking'	fully productive
(c)	*buri/ ppuri*	V/N/ VN/AN	manner of V ing/ being N	*hanasi buri* 'manner of speaking'; *tensai buri* 'being a genius'	semi productive, cf. V forming *bur(u)*
Agent nominals					
(d)	*te*	V	a person who Vs	*hanasi te* 'speaker'; *uri te* 'seller'	productive with agentive V
(e)	*nin* (S J)	V/VN/ AN/N/ root	person	*tutome nin* 'worker'; *binboo nin* 'poor person'; *byoo nin* 'sick person'	semi productive; with free and bound S J forms
(f)	*ka* (S J)	N/VN/ root	profession, person with a habit of Ving	*ongaku ka* 'musician'; *benkyoo ka* 'studious person'; *sak ka* 'writer'	semi productive, refers to an individual property
(g)	*sya* (S J)	VN/A/ N/root	person	*roodoo sya* 'worker'; *zyaku sya* 'the weak'; *zituryoku sya* 'powerful person'	semi productive, often refers to a temporary role
Deadjectival nominals					
(h)	*sa*	A/AN	degree/state of being A	*taka sa* 'height'; *yutaka sa* 'richness'	fully productive
(i)	*mi*	A	property	*ama mi* 'sweet taste'	semi productive
(j)	*me*	A	a certain degree of being A	*taka me* 'somewhat high'	semi productive; with gradable A
Aspectual markers (phrasal affixes)					
(k)	*tyuu* (S J)	VN	while	*[kuruma o unten]*_{VP} *tyuu* 'while driving a car'	Case marking in VP is retained; Genitive *no* on the object NP possible: *kuruma no unten tyuu*
(l)	*go* (S J)	VN	after	*[eki ni tootyaku]*_{VP} *go* 'after arriving at the station'	
(m)	*kake*	V	be about to	*[tegami o kaki]*_{VP} *kake* 'be about to write a letter'	Case marking in VP is retained
(n)	*tate*	V	have just V ed	*[gakkoo ni hairi]*_{VP} *tate* 'have just entered school'	

4.1 Deverbal nominals in Japanese

The important questions concerning nominalization are whether it is to be captured in the lexicon or is a syntactic derivation involving phrases, and if the latter is the case, what level of syntactic projections it involves (see Roeper 2005, among others, for a review). Though lacking a productive affix for forming event-denoting nominals like English *-ing*, *-tion*, and *-al*, Japanese has affixes that attach to verbs and derive nominals that denote manner (*-kata* and *-buri*), agent (*-te*) (Chapter 17 [Ono, this volume]), as well as those that function as aspectual markers (Table 4(k–n)). In this section we consider the affix *-kata* and discuss its syntactic nature.

Morphologically, the suffix *-kata* 'way' attaches to verbal infinitives (*ren'yō* forms) and produces nominals with the meaning 'the manner or way of V-ing' (Table 4(b)). It applies to both simplex and complex or compound verbs with high productivity without any lexical restriction. The nominals suffixed with *-kata* can, as a whole, be used as arguments of a sentence like subject and object, as exemplified by (14).

(14) a. *Sensei   wa   Ken   no   aruki-kata   o   kansatu-si-ta.*
        teacher TOP Ken   GEN walk-way   ACC watch-do-PST
        'The teacher watched how Ken walked.'

  b. *Taizyuu   no   heri-kata   ga   sukuna-i.*
        weight   GEN decrease-way   NOM small-PRS
        'The way of decrease of (my) weight is small.'

In the literature (Sugioka 1986, 1992; Kageyama 1993; Kishimoto 2006), a number of arguments have been constructed to demonstrate the syntactic nature of *-kata* affixation. First, as shown in (14), the nominals derived with *-kata* are compositional and completely transparent in meaning in that they refer to the manner of action or change denoted by the base verb. Their semantic transparency, which presents a sharp contrast to nominals derived by conversion (see below), strongly suggests that *-kata* nominals should not be listed in the lexicon but instead be derived by rule. In terms of semantic transparency, *-kata* nominals thus resemble the so-called "action nominal" in English that are formed with *-ing* (e.g. *[John's careful translating of the speech]$_{NP}$ took many hours*), although they differ in the kinds of meaning they convey, the former denoting a manner of action and the latter an action itself.

The Japanese *-kata* nominalization and the English action nominals are similar in other respects as well. In contrast with V-to-N conversion, which is quite limited in Japanese (see below), *-kata* affixation is extremely productive with no lexical restrictions, thus applying indiscriminately to native infinitive verbs regardless of their morphological complexity (simplex, compound, or derived verbs), as shown by (15).

(15) a. *utai-kata* [sing (simplex verb)-way] 'way of singing'

     b. *kumi-tate-kata* [put-together (compound verb)-way] 'way of assembling'

     c. *tuyo-mari-kata* [become stronger (derived verb)-way] 'way of getting stronger'

In addition, like English action nominals, *-kata* nominals can "inherit" all the arguments and adjuncts of the base verb, realizing them with the genitive case. Observe the correspondence of four noun phrases, *Ken*, *niwa* 'garden', *hana* 'flower', and *mizu* 'water' between the tensed sentence in (16a) and the *-kata* nominalization in (16b).

(16) a. *Ken   ga   niwa   de   hana   ni   mizu   o   yar-u.*
       Ken NOM garden LOC flower DAT water ACC give-PRS
       'Ken gives water to flowers in the garden.'

     b. *Ken   no   niwa   de   no   hana   e   no   mizu   no   yari-kata*
       Ken GEN garden LOC GEN flower to GEN water GEN give-way
       'Ken's way of giving water to flowers in the garden'

Sentential and noun phrase structures, however, differ in the way their arguments and adjuncts are coded. In sentence (16a), the subject, direct object, and indirect object are marked respectively with nominative *ga*, accusative *o*, and dative *ni* whereas the locative adjunct is coded by the postposition *de* (LOC). In the noun phrase in (16b), the genitive particle *no* is added on all of the four arguments, with the concomitant drop of nominative and accusative markers; postpositions like *de* remain as such and are followed by *no*, as in *de no*, and the dative marker *ni* is idiosyncratically replaced by a goal marker *e* 'to, toward' when followed by *no*, as in *hana e no* [flower to GEN]. The *-kata* nominal construction is thus parallel to the English action nominal construction like *Ken's giving of water to flowers in the garden* in manifesting all the arguments and adjuncts of the base verb with appropriate markers.

It has been pointed out that the arguments of the base verb of *-kata* nominals can be suppressed or omitted rather freely: *John no yomi-kata* [John GEN read-way] 'John's way of reading' (Kishimoto 2006: 777). This, however, is not a peculiarity of *-kata* nominals alone but is a general characteristic of Japanese, where arguments recoverable from context are omitted or unexpressed in syntactic structure. *-Kata* nominals also allow an unexpressed subject under an arbitrary interpretation (PRO$_{arb}$), as in the book title *Tikyuu no aruki-kata* [earth GEN walk-way] 'How to walk around the world' or 'Globe-trotter travel guide'. This is possible because nominals need not refer to a specific event. The arbitrary interpretation is available with subjects but not with objects, as shown by the contrast between *susi no tukuri-kata*

[sushi GEN make-way] 'how to make sushi' and *Mari no tukuri-kata [Mari GEN make-way] 'how Mari makes' (N.B. the latter expression is grammatical if the omitted object is not arbitrary but pragmatically inferable).

Kishimoto (2006) argues that the genitive marking on arguments and adjuncts can be accounted for by postulating an underlying syntactic structure like (17a), from which the head verb is raised and attached to the suffix -kata, as in (17b).

(17) a. *[susi o tabe]*$_{VP}$ + -kata b. *[[susi no t] [tabe-kata]*$_N$ *]*$_{NP}$
      sushi ACC eat   way   sushi GEN   eat-way
      'eat sushi'   'manner of eating sushi'

In (17a), the suffix -kata takes scope over VP ('eat sushi'), from which the NP structure in (17b) is derived where the genitive particle is attached on the object and the original accusative marker drops out.

Kishimoto's (2006) syntactic approach should be contrasted with Iida's (2002) lexical approach in the HPSG framework, where [V-kata] as a whole is listed in the lexicon and shares the argument structure with the base verb (e.g. *yar(u)* 'give' [Agent, Goal, Theme] / *yari-kata* [Agent, Goal, Theme]). However, a closer look at relevant data reveals that the case-marking patterns found on some arguments cannot be accounted for by a simple lexical inheritance of the list of thematic roles. Namely, when a dative object is the sole internal argument of a base verb, it is marked with *no*, often without the goal marker *e* that precedes it (Sugioka 1992; Ito and Sugioka 2002: 105).

(18) a. *isu ni suwar-u*   b. *isu no suwari-kata*
          chair DAT sit.on-PRS   chair GEN sit.on-way
          'sit on a chair'   'way of sitting on a chair'

This fact indicates that the dative object here has the status as the internal argument of the base verb rather than an argument theta-marked as goal: if the latter were the case, the object 'chair' would have to be marked with *e no*, but it is highly unnatural: ??*isu e no suwari-kata*. An analogous phenomenon is observed with English action nominals (but not in derived nominals), where a goal object is marked with *of* rather than *into*, as in *entering of the city* as opposed to *entry into/*of the city:* The semantically unmarked preposition *of* corresponds to *no* in (18b) (Ito and Sugioka 2002; Ito 2002).

Even more compelling evidence is available that indicates that -kata nominals are indeed derived in syntax; namely, -kata can productively attach to predicates that are formed by syntactic V-V compounding (19a) or syntactic passive and causative affixation (19b) (see Section 5.2), as well as honorification (19c) (Kageyama 1993; Kishimoto 2006: see also Section 2).

(19) a. *ame   no   huri-hazime-kata*
          rain   GEN   fall-begin-way
          'the way it started to rain'

      b. *Mari   no   hatarak-ase-rare-kata*
          Mari   GEN   work-CAUS-PASS-way
          'the way Mari has been made to work'

      c. *sensei   no   o-hanasi-ni.nari-kata*
          teacher   GEN   HON-speak-HON-way
          'the way the teacher speaks'

Another piece of evidence to the same effect is that *-kata* can attach to VP idioms without cancelling the idiomatic interpretation (Sugioka 1992; Kishimoto 2006).

(20) a. *ocha   o   nigos-u*   b. *ocha   no   nigosi-kata*
          tea   ACC   muddy-PRS   tea   GEN   muddy-way
          'speak evasively (lit. muddy the tea)'   'way of speaking evasively'

The idiomatic meaning of the VP in (20a) is preserved in the *-kata* nominal in (20b), which would not be possible if the scope of the nominalization is just the verb. This is parallel to English action nominals based on some idioms; e.g. *his breaking of the ice between them*.

    Now that we have seen that the behavior of the suffix *-kata* favors the syntactic approach to the lexical one, the question to ask is what level of verbal projection it applies to. First, it has been argued that the scope of affixation does not extend to a higher projection headed by Aspect or Modal. This is because neither a progressive *-te iru* form nor a perfective *-te simau* can be followed by *-kata*: **suwat-te i-kata* [sit-GER be-way] 'the way of being seated', **it-te simai-kata* [go-GER finish-way] 'the way of having gone' (Kageyama 1993), although it is not the case that *-kata* cannot attach to *-te* V complex in general: *(okasi no) [kat-te morai]-kata* [(candy GEN) buy-GER get-way] 'way of having (someone) buy (a candy)'. It has also been observed that the Subject-Object order, which can be scrambled in a sentence, cannot be changed with the nominalized verb, as shown below.

(21) a. *Mari   no   hon   no   yomi-kata*
          Mari   GEN   book   GEN   read-way
          'Mari's way of reading books'

      b. **hon   no   Mari   no   yomi-kata*
          book   GEN   Mari   GEN   read-way

Kishimoto (2006) explains this contrast by saying that the phrasal category to which -*kata* attaches lacks Tense. On the assumption that A-movement requires EPP features, movement of NPs is barred in the nominalization structure.[8]

It should also be noted that -*kata* nominalization shows a remarkable contrast with conversion nominals, i.e. nominals derived from verb infinitives without an overt affix (Table 4(a): see Chapter 9 [Yumoto, this volume]). V-to-N conversion based on simple verbs is not productive in Japanese (Nishio 1961), and the meanings of those that exist are largely unpredictable, as shown by the examples in Table 4(a).[9] Hence these nominals must be listed in the lexicon, and it is predicted that they would not regularly inherit the arguments of the base verb, unlike -*kata* nominals. The prediction is borne out by the contrast between (22b) and (22c) below (Sugioka 1992).

(22) a. *Ken ga   Mari {o   / *ni}   hagemas-u.*
        Ken NOM Mari {ACC / *DAT} encourage-PRS
        'Ken encourages Mari.'

   b. *Ken no   Mari {*no / e no}   hagemasi*
        Ken GEN Mari {*GEN / to GEN} encouragement
        'Ken's encouragement of Mari'

   c. *Ken no   Mari no   hagemasi-kata*
        Ken GEN Mari GEN encourage-way
        'the way Ken encourages Mari'

(22a) shows that the transitive verb *hagemas-u* 'encourage' represents the subject *Ken* with the nominative and the object *Mari* with the accusative. In the nominalized counterpart (22b), the conversion nominal *hagemasi* 'encouragement' calls for the postposition *e* 'to, toward' to realize the "Goal" role of *Mari*. The same pattern can be observed with such conversion nominals as *odosi* 'a threat' and *uyamai* 'respect (N)'. Importantly, although *hagemasi* in (22b) can be interpreted as a result nominal ('words of encouragement'), it can also denote an act of encouragement, as in the sentence *Mari e no hagemasio tuzuke-ru* '(I) continue the encouragement of Mari'.[10]

---

8 Kishimoto (2006) also argues that *kata* attaches to *v*P rather than VP, as evidenced by the projection of the subject NP within the nominalization structure.

9 See Ito and Sugioka (2002: 83 95) for the discussion on how different types of referents denoted by deverbal nouns in Japanese and English can be identified in terms of the base verb's Lexical Conceptual Structure (LCS), as well as Chapter 9 (Yumoto, this volume). An analysis in Distributed Morphology of prefixed deverbal nominals in Japanese can be found in Tagawa (2008), which categorizes roots into Event and State to deal with semantic differences.

10 The nominal *hagemasi* in the result nominal interpretation also appears with an NP followed by the postposition *e* 'to' and the genitive marker (*Mari e no hagemasi* [Mari to GEN encouragement] 'words of encouragement to Mari'), on a par with entity nouns like *okuri mono* 'gift' (*Mari e no okuri mono* [Mari to GEN give thing] 'a gift for Mari').

This shows that *hagemasi* 'encouragement' and other similar conversion nominals denoting an act correspond to Grimshaw's (1990) "simple event nominals" that are lexical in nature. Therefore, *Mari e no* in (22b) is not an argument but a simple modifier (Sugioka 1992; Ito and Sugioka 2002).[11] In contrast, the *-kata* nominal in (22c) does not require the postposition *e* designating the meaning of 'to, toward' because the nominal formed by *-kata* affixation inherits the arguments of the base verb, and the inherited arguments are marked automatically with the genitive marker *no* in noun phrase structure.

From the preceding observations, it is plausible to assume that the affix *-kata* attaches to VP (or *v*P; see note 8) in syntax, whereby the V moves to the position of the nominal containing *-kata* by head-movement (see (17)), thus forming an NP structure. We have also seen that the nominals formed with *-kata* share some properties with action nominals in English, which are arguably also formed in syntax (Ito and Sugioka 2002; Ito 2002 for more evidence).

4.2 Deadjectival nominals

We will now turn to nouns derived from adjectives, with particular focus on two noun-forming suffixes that attach to adjectives: *-sa* (e.g. *omo-sa* 'heaviness' < *omo(i)* 'heavy') and *-mi* (e.g. *omo-mi* 'weight') (see Table 4(h,i)). These two affixes differ appreciably in productivity in that while *-sa* can attach relatively freely to most adjective stems and adjectival nouns, *-mi* is found with only a small subset of frequently used adjectives, with the consequence that there are numerous lexical gaps, such as **kura-mi* 'darkness' as opposed to *akaru-mi* 'light place', and **usu-mi* 'thinness' as opposed to *atu-mi* 'thickness'. The two suffixes also differ in the restriction on the kinds of lexical strata they select: the affix *−mi* is restricted to the bases from the native stratum, whereas *-sa* can freely attach to Sino-Japanese words (*genki-sa* 'cheerfulness') and foreign morphemes (*nau-sa* 'now-ness') as well, and even to newly coined words (*dasa-sa* 'old-fashioned-ness').[12]

Where do these differences in productivity and lexical selection come from? As briefly mentioned in Section 1, Kageyama (1993) argues that the differences can be explained by postulating different levels to which the two affixes attach: stems for *-sa* and roots for *-mi*. This analysis accounts for the observation (Kageyama 1982;

---

11 Compound verbs can form event denoting nominals by conversion more productively than simple verbs (Ito and Sugioka 2002; Chapter 9 [Yumoto, this volume]), and they can appear with the arguments of the base verb, e.g. *hon no kasi dasi* [book GEN lend out] 'lending out of books', which seems to suggest a syntactic derivation. How to account for such cases as opposed to the conversion from simple verbs remains for future study.

12 There are, however, a few sporadic uses of the affix *mi* on the bases of non native strata as well as exceptional gaps in *sa* affixation (see Sugioka 2011).

Sugioka 1986) that *-sa*, but not *-mi*, can attach to compound adjectives as in (23), as well as adjectives derived by various affixes as in (24).

(23)　a.　*ki-yowa-sa* [mind-weak-NMLZ] 'timidity' / **ki-yowa-mi*
　　　　(cf. *yowa-mi* 'weakness')

　　　b.　*koge-kusa-sa* [burn-smelly-NMLZ] 'burned smell' / **koge-kusa-mi*
　　　　(cf. *kusa-mi* 'smell')

(24)　a.　*kodomo-rasi-sa* [child-like-NMLZ] 'childlikeness' / **kodomo-rasi-mi*

　　　b.　*sirooto-ppo-sa* [amateur-ish-NMLZ] 'amateurishness' / **sirooto-ppo-mi*

This contrast can be accounted for by assuming that the lower-level *-mi* affixation cannot follow compounding or other types of affixation that derive stems, while the higher-level *-sa* affixation can (Kageyama 1982, 1993; Sugioka 1986). A similar analysis based on level ordering (see Section 1) has been posited for English nominalization affixes *-ity* and *-ness* (Siegel 1974; Aronoff 1976; Selkirk 1982).

Additional evidence reveals, however, that the difference between the two affixes is not merely one of levels, but one of grammatical components. Namely, the affixation of *-sa* is a syntactic process, while nominals with *-mi* are always listed in the lexicon. (See Note 3 for possibilities of words containing a syntactic affix to be listed in the lexicon.) First, *-sa* can attach to complex adjectival predicates, like "tough" adjectives and desiderative adjectives, that are derived in syntax (Section 3), while *-mi* cannot.

(25)　a.　*kaki-yasu-sa* [write-easy-NMLZ] 'easiness to write' vs. **kaki-yasu-mi*

　　　b.　*asobi-ta-sa* [play-DESI-NMLZ] 'eagerness to play' vs. **asobi-ta-mi*

Second, the affixes differ greatly in their semantic compositionality. The nominals formed with *-sa* are compositional and denote an abstract state or property 'the degree of [A]-ness / the fact of being [A]'; e.g. *naga-sa* 'length, being long'. In contrast, *-mi* nominals can refer to various tangible properties or qualities denoted by the base adjective, so that their actual meanings depend on individual base adjectives: e.g. a form (*maru-mi* 'round shape'), a location (*huka-mi* 'deep point in water'), a taste (*ama-mi* 'sweet taste'), a sensation (*ita-mi* 'pain'), and so on. It is therefore natural to assume that *-mi* nominals are listed in the lexicon along with their idiosyncratic meanings, while *-sa* nominals need not be.

Furthermore, when *-sa* attaches to adjectives that are part of idiom chunks as exemplified in (26), the nominalization preserves the idiomatic meaning. Just as the sentential expression (26a) is ambiguous between literal and idiomatic meanings, so is the *-sa* nominal in (26b). By contrast, the *-mi* nominal in (26c) can only be interpreted literally.

(26) a. *kosi   ga   omo-i.*
     bottom   NOM   heavy-PRS
     'one's bottom is heavy' (literal) / 'be reluctant, be slow to act' (idiomatic)

   b. *kosi   no   omo-sa*
     bottom   GEN   heavy-NMNL
     'the heaviness of one's bottom' (literal) / 'reluctance or slowness to act'
     (idiomatic)

   c. *kosi   no   omo-mi*
     bottom   GEN   heavy-NMLZ
     'heaviness of the bottom' (no idiomatic meaning)

In (26b), the argument of the base adjective must be marked with the genitive *no* in exactly the same way as the arguments in *-kata* nominalization (Section 4.1). This shows that the suffix *-sa* has the whole idiom as its scope, and that the adjective stem (*omo-* 'heavy') is attached to the suffix by head-movement, as in *[[kosi no t] [omo-sa]$_N$]$_{NP}$*. The parallelism can also be seen in the case of adjectives taking two nominative NPs, which are both turned into genitive in the nominalization structure: *Ken ga ryoori ga uma-i* [Ken NOM cooking NOM good.at-PRS] 'Ken is good at cooking' and *Ken no ryoori no uma-sa* [Ken GEN cooking GEN good.at-NMNL] 'Ken's being good at cooking'.

    One notable exception with regard to the case marking of the argument NP can be found in the adverbial construction with *-sa ni* [-*sa* DAT] or *-sa no amari* [-*sa* GEN excess] 'out of desire to', where the affixation of *-sa* retains the clausal case markings on arguments (i.e. nominative, accusative, or dative):[13]

(27) *Kane   **ga**   hosi-sa   ni   uso   o   tuk-u.*
     money   NOM   want-NMLZ   DAT   lie   ACC   tell-PRS
     '(I) lie out of wanting money.'

   cf. **kane   no   hosi-sa   ni*
      money   GEN   want-NMLZ   DAT

In this particular construction, it is plausible to assume that the suffix + Dative particle complex (*-sa ni*) select the AP headed by *hosi-* 'want' as in [*kane ga hosi*]$_{AP}$ *-sa ni* in syntactic structure (Sugioka 1986; Kageyama 1993), and the suffix *-sa* gets attached to the head adjective at the phonological structure, thereby maintaining

---

13 The *sa ni* construction is mostly limited to an A or AN denoting desire or lack thereof; hence, the same case marking pattern is observed for the construction formed with AN *iya* 'do not want': *benkyoo ga iya (da)* [study NOM dislike (COP)] 'do not want to study' / *benkyoo ga iya sa ni* [study NOM dislike NMLZ from] 'because of reluctance to study'. Furthermore, the construction can also involve an AP headed by the desiderative *ta(i)*, a phrasal affix taking a VP with the accusative or nominative marking on the object NP of the verb (Section 3.3):

[[Oisii   raamen   o/ga   tabe]$_{VP}$ ta] sa   ni   gyooretu si ta.
delicious noodle   ACC/NOM   eat DESI NMLZ DAT   queue do PST
'(We) lined up in a queue out of desire to eat delicious noodles.'

the case marking within the AP structure.[14] This happens only in the special constructions formed by *-sa ni* or *-sa no amari*, and nominalization by *-sa* otherwise forms an NP structure as we saw above.

To recapitulate, in Japanese deverbal and deadjectival nominalization there are both fully productive syntactic processes (*-kata* affixation and *-sa* affixation) and less productive lexical processes involving idiosyncrasies that call for listing in the lexicon (V-to-N conversion and *-mi* affixation). The affixes *-sa* and *-mi* are especially noteworthy in providing minimal pairs (*yowa-sa* / *yowa-mi* 'weakness') with contrasting behaviors for experimental studies (see Section 6). These contrasting behaviors of nominalization affixes attest to the need for recognizing the variety of nominalization processes with regard to the syntax-lexicon dichotomy.

5 Verb formation

This section surveys verb-forming suffixes, dividing them into lexical and syntactic suffixes, as exemplified in Table 5.

Table 5: Verb-forming affixes in Japanese

	suffix	base cat.	meaning	examples	comments
Lexical affixes					
(a)	*r(u)*	N/A/ mimetic root	Verbalizer	*memo r(u)* 'take a memo'; *huto r(u)* 'become fat'; *teka r(u)* 'glitter'	unproductive, used to coin a new verb in an informal style
(b)	*mar(u)*	A	inchoative	*taka mar(u)* 'become higher in degree'	unproductive, intransitive: cf. *me(ru)*
(c)	*me(ru)*	A	causative	*taka me(ru)* 'make x higher in degree'	unproductive, transitive: cf. *mar(u)*
(d)	*m(u)*	A	to feel A	*kanasi m(u)* 'feel sad'	unproductive
(e)	*bur(u)*	N/A/AN	act in a manner of N/A	*otona bur(u)* 'act like an adult'; *ara bur(u)* 'act violently'	semi productive: cf. N forming *buri*
(f)	*mek(u)*	N/A	to bear/show the property of N/A a little	*haru mek(u)* 'be more like spring'; *huru mek(u)* 'look old'	unproductive, archaic
Syntactic affixes					
(g)	*gar(u)*	A/AN	express the feeling of A	*kanasi gar(u)* 'express sadness'; *husigi gar(u)* 'wonder'	productive with subjective A; used with 3rd person subject NP
(h)	*(s)ase(ru)*	V	causative	*ik ase(ru)* 'make sb go'	fully productive, syntactic embedding

---

14 The nominalizing suffix *sa* in this construction constitutes part of the complex (*sa ni* or *sa no amari*) that can be perceived as a clitic, with the function of introducing a clause that denotes a reason ('because of (the desire to V)'). See Chapter 14 (Kageyama, this volume), where the aspectual bound morphemes retaining the sentential case marking (see (2a) and Table 4(k n) above for examples) are analyzed as clitics.

5.1 Lexical verb formation

Apart from addition of the verb *suru* that makes up a syntactic light verb construction, lexical derivation of verbs from other categories is limited in contemporary Japanese; inflected verbs cannot be freely coined in the native Japanese stratum. As noted in Chapter 1 (Kageyama and Saito, this volume), the suffix *-r* (see Table 5(a)) is marginally available in colloquial style to derive a verb from a noun, as in *guti-r(u)* 'to complain' from *guti* 'complaint'. The process is not productive, and most new formations like *bagu-r(u)* '(for a PC) to malfunction' from *bagu* 'bug' are creatively formed coinages and are used only in casual language, except for a very few established words like *kumo-r(u)* 'become cloudy' from *kumo* 'cloud' and *kage-r(u)* 'darken (intr)' from *kage* 'shadow'. The same verb-forming suffix *-r* can also attach to an adjectival root, as in *yowa-r(u)* 'weaken (intr.)' from *yowa(i)* 'weak' and *nibu-r(u)* 'become dull' from *nibu(i)* 'dull', but this usage is restricted to only a few established words, and is even more limited in productivity.

It should be noticed that the semantic interpretation of verbs derived by the suffix *-r* is irregular, exhibiting a variety of semantic types as shown in (28):[15]

(28) a. activity verbs: *kyodo-r(u)* 'act strangely' < *kyodoo-husin* 'suspicious behavior', *gugu-r(u)* 'search by Google' < *guuguru* 'Google'

   b. change-of-state verbs: *debu-r(u)* 'become fat' < *debu* 'a fat person', *paniku-r(u)* 'become panicky' < *panikku* 'panic'

   c. verbs of transportation: *taku-r(u)* 'go by taxi' < *takusii* 'taxi', *tyari-r(u)* 'go by bike' < *tyari* 'bike'

   d. verbs of utterance: *yazi-r(u)* 'heckle' < *yazi* 'heckling', *dabe-r(u)* 'chat' < *daben* 'idle talk'

This is similar to cases of English N-to-V conversion and verb-forming affixation (with suffixes like *-en, -ize, -ify*), which exhibit a large variety in meanings (Clark and Clark 1979; Lieber 1998; Plag 1999). The unpredictability in meaning shown in (28) lends support to the lexical status of the suffixation.

As exemplified by (29), *-mar(u)* and *-me(ru)* (Table 5(b,c)) look like de-adjectival verb-forming affixes that select for stage-level adjectives to derive change-of-state verbs.

(29) *taka-mar(u)* (intr.) / *taka-me(ru)* (tr.) 'heighten', *tuyo-mar(u)* (intr.) / *tuyo-me(ru)* (tr.) 'strengthen', *hiro-mar(u)* (intr.) / *hiro-me(ru)* (tr.) 'widen'

In present-day Japanese, these suffixes are lexically restricted and are not very productive, exhibiting many lexical gaps (e.g. ***asa-mar(u)** (intr.)/***asa-me(ru)** (tr.) <

---

15 As can be seen in (28), *r* suffixation sometimes involves clipping, especially when the base is a borrowing with many moras.

asa(i) 'shallow' as contrasted with *huka-mar(u)* (intr.) / *huka-me(ru)* (tr.) 'deepen' < *huka(i)* 'deep') and some irregular phonological changes (e.g. *se̱ba-mar(u)* 'get narrow' <*se̱ma(i)* 'narrow'). In addition, these affixes select for a root, so that they cannot attach to a polymorphemic stem, as illustrated by **oku-huka-me(ru)* <*oku-huka(i)* (end-deep) 'profound'.

Historically, *-mar(u)* and *-me(ru)* come from a verb-forming suffix *-m* followed by *-ar* and *−e*, respectively. In fact, *-m* is detected with subjective adjectives (Table 5(d)) as in the intransitive *ita-m(u)* 'hurt' (from *ita(i)* 'painful'), which alternates with the transitive *ita-m-e(ru)* 'hurt'. On the other hand, *-ar* and *-e* are not limited to de-adjectival verbs but are involved with the transitivity alternations in a wide range of verbs, as in (30).

(30)   *butuk-ar(u)* (intr.) / *butuk-e(ru)* (tr.) 'hit', *sag-ar(u)* (intr.) / *sag-e(ru)* (tr.) 'lower', *tunag-ar(u)* (intr.) / *tunag-e(ru)* (tr.) 'connect'

In the research on transitive-intransitive alternations of Japanese verbs, it has been controversial which of the transitive *-e* and the intransitive *-ar* in such pairs is basic or whether they are both basic (equipollent). The problem is in fact even more complex in that the same suffix *-e* manifests itself in intransitive verbs like those in (31).

(31)   a.   *tok-e(ru)* (intr.) / *tok-as(u)* (tr.) 'melt', *sam-e(ru)* (intr.) / *sam-as(u)* (tr.) 'cool'

   b.   *sak-e(ru)* (intr.) / *sak(u)* (tr.) 'split', *or-e(ru)* (intr.) / *or(u)* (tr.) 'break'

These examples show that the relationships between morphological makeup, meaning, and transitivity are not fully systematic or predictable in present-day Japanese,[16] and hence the verb alternations should be considered lexical, in stark contrast to the fully regular causative suffix *-sase*, which is syntactic (see Section 5.2).

The semantics of verbs derived with the suffixes *-me/-mar* also exhibits irregularity to some extent: these suffixes derive change-of-state verbs, but a closer look reveals that they comprise two different types of intransitive-transitive pairs as shown in (32)–(33) (Sugioka 2002):

(32)   a.   *Heya   ga   atata-mar-u.*
            room   NOM   warm-VBLZ-PRS
            'The room warms.'

   b.   *Ken   ga   heya   o   atata-me-ru.*
            Ken   NOM   room   ACC   warm-VBLZ-PRS
            'Ken warms the room'

---

16 There is abundant literature on the transitivity alternations of Japanese verbs involving not just these two suffixes but a complex array of other suffixes. See Jacobsen 1992, forthcoming; Suga and Hayatsu (eds.) 1995; Kageyama 1996; Matsumoto forthcoming; Narrog forthcoming; for representative approaches to the issues.

(33) a. *Kaze   ga   tuyo-mar-u.*
         wind   NOM   strong-VBLZ-PRS
         'The wind gets strong'

      b. *Kaze   ga   ikioi   o   tuyo-me-ru.*
         wind   NOM   power   ACC   strong-VBLZ-PRS
         'The power of the wind gets strong'

      c. **Ken   ga   kaze   o   tuyo-me-ru.*
         Ken   NOM   wind   ACC   strong-VBLZ-PRS
         'Ken makes the wind strong.'

In the pairs like (32), intransitive verbs denote some change of state forced by an external instigator as expressed by the transitive counterparts, exhibiting a case of de-causativization in Kageyama's (1996) terms. In pairs like (33), on the other hand, intransitive verbs denote internally caused change-of-state, where no external instigator is involved. Thus the transitive counterpart like (33c) is unacceptable, while transitive -*me(ru)* verbs can be used in sentences like (33b), where the entity undergoing the change of state is expressed as the subject, rather than as the object. This semantic variation is not totally predictable, which gives support to the lexical status of these suffixes.

The observations in this section, taken together, point to the lexical status of the suffixes -*r* and -*me/mar*. A syntactic approach by Hale and Keyser (1993) analyzes denominal verbs as involving a VP structure in which the base noun is incorporated into the verb position to form a verb. In view of the idiosyncrasy displayed by verbalization processes in Japanese as well as in English, however, the syntactic approach seems unable to wholly capture the characteristics of the processes (Kiparsky 1997; Kageyama 1997).

5.2 Syntactic verb formation

In contrast to the lexical verb formation discussed in Section 5.1, Japanese has a couple of verb-forming affixes that are arguably syntactic. This subsection deals with two of them, the suffix -*gar* 'express (one's feeling)', which derives a transitive (or unergative) verb from an A or an AN, and the suffix –*sase*, which derives a causative verb from a verb.

As shown in (34), -*gar* 'express (one's feeling)' attaches to an A/AN to derive a verb, where the subject of the base A/AN is changed to the object of the derived verb:

(34) a. *Kono geemu ga omosiro-i / taikutu da.*
　　　 this　game　NOM　enjoyable-PRS / boring COP
　　　 'This game is fun / boring.'

　　 b. *Kodomo ga geemu o omosiro-gar-u / taikutu-gar-u.*
　　　 child　　NOM game　ACC enjoyable-express-PRS / boring-express-PRS
　　　 'Children express enjoyment/boredom of games'

This suffix, though its base is exclusively limited to an A/AN that denotes subjective feelings and sensations, is quite productive within that semantic domain. Moreover, it is often attached to complex adjectives that already involve syntactic suffixes like *-ta(i)* 'want to' or *-niku(i)* 'difficult to' (see Section 3.3), which confirms that *-gar* suffixation is also syntactic.

(35) a. *kodomo ni hon o yom-ase-ta-gar-u*
　　　 child　　DAT book ACC read-CAUS-DESI-express-PRS
　　　 'express the desire to make the children read a book'

　　 b. *aruki-niku-gar-u*
　　　 walk-tough-express-PRS
　　　 'express the difficulty to walk'

　　It is interesting to note that the syntactic suffix *-gar* (Table 5(g)) and the lexical suffix *-m* (Table 5(d)) both select for adjectival bases of the same semantic type, namely those denoting subjective feelings. Hence we find doublets like: *ayasi-gar(u) / ayasi-m(u) < ayasi(i)* 'suspicious', *natukasi-gar(u) / natukasi-m(u) < natukasi(i)* 'nostalgic', *kurusi-gar(u) / kurusi-m(u) < kurusi(i)* 'painful'. As expected, however, a number of lexical gaps are found for the lexical suffix *-m*, but not for the syntactic suffix *-gar* in the semantic domain: *omosiro-gar(u) / *omosiro-m(u) < omosiro(i)* 'enjoyable', *kowa-gar(u) / *kowa-m(u) < kowa(i)* 'scary'. Also as predicted, the lexical suffix *-m* is incapable of appearing after syntactically derived adjectives as in **yom-ase-ta-m(u)* [read-CAUSE-DESI-suffix].

　　Next, let us turn to a typical clause-embedding suffix *-(s)ase* 'cause', which constitutes one of the most well-discussed topics in the field of Japanese syntax (see Miyagawa 1999 and Harley 2008 for literature review). One of the striking peculiarities of this suffix is that it is capable of making a causative verb out of an intransitive verb even though there may be a lexically derived transitive verb that represents a causative meaning of the same intransitive verb. If we refer to a causative verb with *-(s)ase* as a "*sase*-causative" (abbreviated to SC) and to a transitive change-of-state verb as a "lexical causative" (LC), we find minimal pairs of an LC (e.g. *narab-e* [align-TR] 'put in line' < intr. *narab-* 'stand in line') and an SC (e.g. *narab-ase* [align-CAUS] 'make someone stand in line'), as in (36).

(36) a. *Sensei   ga   hon   o   tana   ni   narab-e-ta.* [LC]
        teacher   NOM   book   ACC   shelf   LOC   align-TR-PST
        'The teacher arranged books on the shelf.'

     b. *Sensei   ga   seito   o   kooteil   ni   narab-ase-ta.* [SC]
        teacher   NOM   pupil   ACC   school.yard   LOC   align-CAUS-PST
        'The teacher made the pupils align in the school yard.

The two verbs can co-exist because they represent different kinds of causative meaning: LC denotes manipulative causation, where the causer directly manipulates the causee (the entity denoted by the direct object) and causes the change of state, whereas SC denotes directive causation, where the causer gives direction to the causee to perform the action denoted by the verb stem. Thus the causee argument of an LC typically denotes an inanimate entity while that of an SC denotes an animate (typically human) entity that can perform some action.

In addition to this semantic difference, it is known that the two types of causatives exhibit different syntactic behavior. Although SC and LC sentences share the case marking patterns of mono-clausal sentences, as shown in (36), SC sentences display some properties suggesting bi-clausal syntactic structure,[17] while LC sentences never show such properties. The most convincing evidence is obtained from *zibun*-binding. Since *zibun* 'self' has subject-orientation in the selection of its antecedent and allows long-distance binding, *zibun* in a single clause is never ambiguous (the only possible antecedent being the subject), while that in an embedded clause is ambiguous (with both matrix subject and the embedded subject being potential antecedents). With these properties of *zibun* in mind, let us consider the following SC and LC sentences with *zibun* (Shibatani 1976):

(37) a. *Mari$_i$ ga   musuko$_j$ o   zibun$_{i/j}$ no   zitensya ni   nor-ase-ta.* [SC]
        Mari   NOM   son   ACC   self   GEN   bicycle   LOC   ride-CAUS-PST
        'Mari made her son ride her/his own bicycle.'

     b. *Mari$_i$ ga   musuko$_j$ o   zibun$_{i/*j}$ no   zitensya ni   nose-ta.* [LC]
        Mari   NOM   son   ACC   self   GEN   bicycle   LOC   lift-PST
        'Mari lifted her son onto her/*his own bicycle.'

As indicated by the subscripts, (37a) is ambiguous with either Mari or her son being the possessor of the bike, while (37b) does not exhibit the ambiguity. This

---

17 It should be noted, however, that SC sentences exhibit some mono clausal properties including the licensing of the negative polarity item *sika* 'except' as well as the case marking patterns, supporting the mono clausal analyses as proposed by Miyagawa (1989) in the Government and Binding (GB) tradition, Manning, Sag and Iida (1999) in HPSG, and Matsumoto (1996) in Lexical Functional Grammar (LFG).

strongly suggests that an SC sentence like (37a) has a bi-clausal structure where the accusative-marked NP acts as the subject of the embedded clause, and hence as the antecedent of *zibun*, while (37b) has a mono-clausal structure.

Another difference between SC and LC is that the formation of the former is a regular and productive process while that of the latter is not. First, the suffix *-(s)ase* can be attached to any verb as far as its semantic restriction is satisfied: as noted above, since SC denotes directive causation, the base verb must denote some agentive action controllable by the causee,[18] but otherwise *-(s)ase* suffixation is fully productive. In contrast, LC formation has lexical gaps: for instance, we do not have lexical causative verbs corresponding to *sabi-ru* 'corrode' (*sabi-s-u*) or *tukare-ru* 'get tired' (*tukar-as-u*). Second, the morphological forms of SC are totally regular: the suffix *-sase* is attached to vowel-ending verb stems (e.g., *tabe-ru > tabe-sase-ru*) while the allomorphic variant *-ase* is attached to consonant-ending verb stems (e.g., *aruk-u > aruk-ase-ru*). In contrast, the forms of LC verbs are quite complicated and not necessarily predictable as discussed in Section 5.1.

In summary, SC and LC exhibit different properties in terms of regularity, productivity, and syntactic structure as well as semantic interpretation. Above all, if the syntactic analysis postulating bi-clausal structure for SC is on the right track, the SC formation (i.e., *-(s)ase* suffixation) can best be treated as a syntactic process, which is also compatible with its full regularity and high productivity. In contrast, LC does not exhibit any of these characteristics, and LC verbs behave like ordinary transitive verbs. Thus the two causative constructions provide ideal material for investigating the validity of the modular approach to word formation.

6 Evidence from neurolinguistic experiments

So far we have seen various linguistic phenomena that can best be analyzed in the modular/hybrid approach. It is true that in phenomena like nominalization and causative verb formation, a pair of word formation processes with similar functions exhibit different properties, thus supporting the deployment of two different components to deal with them. Nevertheless, both syntactic and lexicalist approaches have developed mechanisms to deal with the relevant differences, and hence it is very difficult to decide, drawing solely on linguistic evidence, whether or not the hybrid approach is the *only* appropriate framework (see Section 7). In this context, evidence from psycho- or neurolinguistics can potentially play a vital role: If two word-formation processes are found to employ different mental or neural mechanisms, then it is strongly suggested that the two processes are different in nature,

---

18 This is not the case for "adversity" causative construction, which is also formed with the suffix *(s)ase*: *yasai o kusar ase ru* [vegetable ACC rot CAUS PRS] 'let vegetable rot' (Harley 2008).

not just in the degree of regularity and productivity, thus providing support for the hybrid approach.

One of the major controversies on morphology in the field of neurolinguistics is, as mentioned in Section 1, the so-called "past-tense debate" in English (Pinker and Ullman 2002). Proponents of the dual mechanism model (Pinker 1999) argue that some regular and productive morphological processes like regular inflections in English involve rule-based computation while other less regular and less productive processes like irregular inflections in English involve associative memory. In this model, the regular past tense formation like *walk > walked*, for instance, is processed by application of a symbolic rule of *-ed* suffixation, while irregular past tense formation like *sing > sang* or *blow > blew* is dealt with by associative memory, which is based on lexical memory with a network-like structure, i.e. patterns based on phonological similarities, like the *Xing/Xang* pattern (*ring/rang, spring/sprang* etc.) or the *Xow/Xew* pattern (*throw/threw, grow/grew* etc.). Although overapplication of the past-tense formation in child language and experimental situations is observed for both regular and irregular inflections (e.g. *holded* instead of *held*, *glew* instead of *glowed*), the overregularized form (*holded*) is ascribed to the application of rule, while the "over-irregularization" (*glew*) can be attributed to an analogy based on associative memory. Opposed to this view is the single mechanism model (Joanisse and Seidenberg 1999; Bybee 1995), which maintains that one and the same mechanism can deal with both regular and irregular processes. Evidence from neurolinguistic experiments has been accumulating in favor of the dual mechanism model, as surveyed in Ullman (2001). Although most of the studies are devoted to European languages like English and German, some studies on Japanese derivational morphology have been reported, which will be reviewed in what follows.

6.1 Evidence from Japanese nominalizing suffixes *-sa/-mi*

As discussed in Section 4.2, the nominalizing suffixes *-sa* and *-mi* constitute an ideal pair of word derivations with similar functions but contrastive productivity and regularity. Whereas *-sa* suffixation is fully regular and productive, *-mi* suffixation has lexical gaps as well as semantic irregularities, suggesting a syntax-lexicon dichotomy in word formation. Hagiwara et al. (1999) exploited this pair of suffixes to test the validity of the dual mechanism model in Japanese derivational morphology, conducting experiments on two different types of aphasic patients.

Aphasia caused by damage in different regions of the brain are known to exhibit different types of atypical language use (Hagiwara 1998). For the discussion in this section, two types of aphasia are relevant: Broca's aphasia and *gogi* (word-meaning) aphasia. The left anterior perisylvian tissue, known as Broca's area, and its adjacent areas are associated with Broca's aphasia. Patients with damage in these areas often exhibit linguistic behavior known as agrammatism, where they have difficulty in

processing function words like articles or prepositions, as well as grammatical morphemes like past tense or plural morphemes. In Japanese, omission and/or misuse of case particles is characteristic of the utterances of patients with Broca's aphasia. Experimental studies have also shown that patients have difficulty in understanding passives, cleft sentences, scrambled sentences, and other movement-related constructions (Grodzinski 1986). The exact nature of agrammatism is still in debate, but it can safely be assumed that the deficit is related to syntactic computation.

Gogi aphasia, on the other hand, is associated with damage in the left-middle and inferior temporal areas, and patients characteristically show selective loss of word meanings, in contrast to relatively well-retained syntax. They often have difficulty in understanding the meaning of fixed figurative expressions like proverbs (Hagiwara 1998). Thus it can be assumed that *gogi* aphasics' problems lie in the memory system of the lexicon.

Hagiwara et al. (1999) devised two types of contexts for nominalization, taking advantage of semantic differences between the two suffixes as discussed in Section 4.2: one context in which -*sa* forms sound more natural (*sa*-discourse) and the other in which -*mi* forms are more natural (*mi*-discourse). A nonce adjective created for the purpose of this experiment was given in each context, and the participants including Broca's aphasics and *gogi* aphasics as well as non-aphasic controls were asked to select either the -*sa* form or the -*mi* form. The results revealed a sharp contrast between the two types of aphasics. In short, Broca's aphasics selected -*sa* forms only 50% of the time (i.e. at the chance level) even in *sa*-discourse, while *gogi* aphasics tended to select -*sa* forms even in *mi*-discourse. This dissociation strongly suggests that the two suffixation processes are subserved by different brain areas, and hence are arguably different in nature.

It should be noted that the syntax-lexicon dichotomy revealed in this study involves the same type of word formation, namely affixation. Previous studies in the dual mechanism model have centered on inflectional processes of European languages like English and German, where regular inflections tend to take the form of suffixation while irregular ones often involve non-affixal morphology (e.g. vowel change in *sing/sang*), and hence the regular-irregular difference could be relegated to phonological factors (Johanness and Seidenberg 1999). Such an account is untenable for the Japanese data, pointing to the significance of studying typologically different languages like Japanese with agglutinative morphology.

6.2 Evidence from Japanese causative verb formation

As surveyed in Section 5.2, *sase*-causatives (SCs) and lexical causatives (LCs) constitute another pair appropriate for testing the syntax-lexicon dichotomy of word formation. An SC verb exhibits bi-clausal properties even though it is morphologically one word, as well as showing full productivity and regularity, and hence syntactic

analysis of -(s)ase suffixation has been proposed. LC verbs, in contrast, exhibit the syntactic properties of a simple transitive verb, and hence the lexical analysis. This subsection reviews three experimental studies exploiting the pair.

First, Sugioka, Ito and Hagiwara (2001) examined whether or not Broca's aphasics show different levels of performance on SC and LC production and comprehension. Three types of stimulus sentences were constructed: one in which an intransitive verb is appropriate, one in which an LC verb is the best fit, and one in which an SC verb is well-formed. Using these sentences, two experiments were conducted: one was a fill-in-the-blank type production task, and the other was a forced-choice task. Patients of aphasia showing agrammatic comprehension patterns and non-aphasic controls participated in both of the experiments, with an appropriate interval of time. The results showed that the agrammatic patients obviously had problems in the use of SC verbs: in both the production task and the forced-choice task, their performance was much worse for SC verbs than for LC and intransitive verbs. A picture-matching experiment was also conducted to test the participants' comprehension of thematic roles in SC and LC sentences. It was revealed that the patients' performance on LC sentences was significantly better than on SC sentences, providing another piece of evidence that SC sentences are more difficult to process than LC sentences for agrammatic aphasics. Given the characteristics of agrammatism surveyed in Section 6.1, a plausible conclusion to be drawn from these results is that the processing of SC verbs involves some extra syntactic computation which is not employed in the processing of LC verbs and intransitives.

Another experiment was conducted by Fukuda and Fukuda (2001) on children with Specific Language Impairment (SLI). SLI is a type of developmental impairment, which is known to be genetic. Children with SLI show difficulty in language use with no apparent other disorders like hearing loss or general developmental delay. The most prominent weakness of SLI children lies in nonlocal, hierarchical, or combinatorial aspects of language, including regular inflection and nonlocal dependencies like *wh*-questions, while other aspects, including lexical and semantic representations, are comparably unimpaired (van der Lely and Pinker 2014). Thus we may safely assume that SLI children have problems in grammatical computation, a situation similar to agrammatic aphasic patients. Fukuda and Fukuda (2001) conducted a production experiment in which SLI children as well as typically developing controls were asked to fill in the verb-ending forms of LC verbs and SC verbs. The results showed that SLI children had problems producing the SC verbs in contrast to their near-typical ability to produce LC verbs.

The third study was conducted by Ito, Sugioka and Hagiwara (2009), using Event Related Potential (ERP) measurement techniques. ERPs are scalp-recorded changes in electrical activity of the brain, triggered by certain sensory, cognitive, or motor events, and their measurement offers a powerful tool for investigating real-time language processing, with different ERP components reflecting different aspects

of language comprehension (Kutas and Federmeier 2009). When the participants (native Japanese speakers) read unacceptable LC sentences, an ERP component known as N400, which is robustly elicited in relation to semantic violations, was observed. On the other hand, when they read unacceptable SC sentences, an ERP component known as Left Anterior Negativity (LAN), which has been argued to be related to syntactic violations, was observed followed by an N400. The difference in the observed ERP components arguably reflects the difference in the nature of the processing of the two types of causative constructions, namely, only SC sentences involve syntactic bi-clausal structures.

In this section, we have reviewed some neurolinguistic studies on Japanese word-level processing that provide support for the dual mechanism view. The aphasia studies, based on the idea of localization of brain functions, revealed that two different types of word formation are subserved by different brain areas, which strongly suggests that the two types involve different mental and neurological mechanisms. In addition, the study on SLI children, as well as the ERP study, provide evidence for the difference in mental and neurological mechanisms involved in word formation processes from different dimensions, namely, language use by a genetically impaired population and the brain function of "normal" population during on-line processing. Hence the syntax-lexicon dichotomy is supported by experimental studies.

7 Conclusion and future research perspectives

In this chapter we have observed various processes of derivational morphology in Japanese and discussed their characteristics that suggest the dual nature of word formation. We have also surveyed a series of neurolinguistic experiments that give support to the syntax-lexicon dichotomy.

Examination of a wide range of affixation in Sections 2 to 5 confirmed some correlations among the fundamental properties of word formation processes: productivity, semantic compositionality, and various levels at which affixation applies. For example, those suffixes that attach to roots, like the nominalizing suffix -*mi* and verb-forming -*r*, exhibit idiosyncratic meaning and low productivity. Those attaching to stems or words, like -*sa* and -*gar*, on the other hand, produce words with compositional meaning, and are fully productive. Note in this connection that the suffix -*gar* might seem only semi-productive in view of the small number of the adjectives it attaches to. However, when its semantic restriction (i.e. selecting for adjectives of emotion or senses) is taken into consideration, the suffix turns out to be fully productive (e.g. *uzatta-gar(u)* 'express irritation' (colloq.)).

Furthermore, among productive affixes, some can be attached to a unit larger than word, namely word[+] and phrase in syntax, which reflects the agglutinative

nature of Japanese. More specifically, we have observed that in the case of some syntactic suffixes (e.g. *-kata* and *-sa*) the base is attached to the affix by head-movement, while there are cases of phrasal affixation (e.g. *-ta(i)* and *-yasu(i)*) that can leave the case marking within the VP intact (Section 3.3) as well as clause-embedding suffixation (e.g. *-sase* in Section 5.2). It is also noteworthy that syntactic affixes like passive *-rare* and clause-embedding causative *-sase* can be followed by category-changing derivational suffixes like adjective-forming *-ta(i)*, verb-forming *-gar*, and nominalizing *-sa* as in *hatarak-ase-rare-ta-gari-yasu-sa* (work-CAUS-PASS-DESI-express-easy-NMLZ) 'likelihood of expressing desire to be forced to work'. The presence of affixes like these calls for an analysis that captures the way in which derivational affixation is closely intertwined with syntax.

Finally, also worth mentioning is the fact that there are noun-forming affixes that inherently take phrases and function as aspectual markers (Table 4(k-n)), as well as adjective-forming affixes which, in addition to applying at the word level, may attach to a clause to express modality (evidentiality) and function as auxiliaries (Table 3 (c,d,g)). These aspectual markers and auxiliaries are presumably attached in phonological structure.

Based on these observations on Japanese derivational affixes, a hybrid or modular approach seems to be a natural way to account for their wide range of behaviors. Namely, idiosyncrasy and unproductivity can be appropriately relegated to the lexicon on one hand, and regularity, compositionality, and productivity to syntax on the other. This also provides a way to deal with the systematic correlation between these characteristics and the level of application.

As mentioned in Section 1, however, there have been opposing views proposed on the place of morphology in grammar, namely that word formation can and should be dealt with in one component, either in syntax or in the lexicon. In these theoretical frameworks, intricate systems have been devised in order to deal with these contrasting behaviors of derivational affixes with regard to regularity and idiosyncrasy. In syntactic approaches difference in height in the syntactic structure is claimed to offer a key to account for the dichotomy. In Distributed Morphology, for example, affixes attaching to a word, which consists of a root ($\sqrt{}$) and a category node (n, a, v), show regularity, while those attaching to a root alone (an atomic element devoid of category) exhibit idiosyncrasy and unproductivity (Marantz 2007; Harley 2008). Lexicalist approaches like HPSG (Pollard and Sag 1994) and Construction Morphology (Booij 2010), on the other hand, capture the regularity and idiosyncrasy in terms of the hierarchically organized lexicon: morpheme/word-specific idiosyncrasies can be captured at the bottom of the hierarchy, while the degree of regularity corresponds to different levels of abstractions in the schemas at higher levels.

Consequently, it has become extremely difficult to decide which approach should be taken based solely on linguistic facts. One possible way out of this

problem is to regard the issue as a question of whether the difference in regularity/ idiosyncrasy is a reflection of the difference in the nature of the mental processes involved, which would allow findings from a neurolinguistic perspective to offer new insight: if we can claim that two word formation processes employ different neurolinguistic mechanisms, then it would be very hard to trivialize the difference and ascribe both to one grammatical component. To explore this possibility, collaboration between theoretical linguistics and neurolinguistics is rapidly developing, and studies on Japanese word formation in this area, though relatively scarce, have made substantial contributions as surveyed in Section 6. More studies in this field are expected to provide new evidence not obtainable through theoretical investigations based on linguistic facts alone, offering a new perspective to the workings of word formation processes.

The syntax-lexicon dichotomy constitutes a long-standing major issue in theoretical linguistics. As discussed in Sections 2 to 5, Japanese provides a rich source for examining the correlations among the diverse properties displayed by various affixes in order to elucidate the relevant factors for studying this important issue. There are a number of problems concerning affixation and the nature of the lexicon that remain for future study. In particular, whether affixes have lexical entries (i.e. are listed in the lexicon) as words do, and if so, what representations they should be given are fundamental issues that call for more investigation on empirical grounds, not only from a theoretical viewpoint. These inquiries lead to the question of the division between affixation and compounding, an especially important but thorny question for an agglutinative language like Japanese. Moreover, the wide range of productivity in the processes of derivational affixation surveyed in this chapter point to the necessity of looking further into the different types of restrictions imposed on each process and of providing a theoretical model for their accounts. Further investigation into these issues will provide a better understanding of the nature of word-formation processes, namely, the mechanisms that underlie the very basic unit of language.

Acknowledgments

References

Anderson, Stephen R. 1992. *A morphous morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Bybee, Joan L. 1995. Regular morphology and the lexicon, *Language and Cognitive Processes* 10(5). 425–55.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs and Peter Rosenbaum (eds.), *Readings in English transformational grammar*. 184–221. Waltham, MA: Ginn.

Clark, Eve V. and Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55. 767–811.

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Farmer, Ann. K. 1980. *On the interaction of morphology and syntax*. Cambridge, MA: MIT dissertation.

Fukuda, Shinji and Suzy E. Fukuda. 2001. The acquisition of complex predicates in Japanese specifically language-impaired and normally developing children. *Brain and Language* 77. 305–320.

Fukushima, Kazuhiko. 2014. Morpho-semantic bracketing paradox and compositionality: The implications of the sized inalienable possession construction in Japanese. *Journal of Linguistics* 50(1). 91–141.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Grodzinski, Yosef. 1986. Language deficits and the theory of syntax. *Brain and Language* 27. 135–159.

Hagiwara, Hiroko. 1998. *Nō ni idomu gengogaku*. [Linguistics challenging the brain.] Tokyo: Iwanami Shoten.

Hagiwara, Hiroko, Yoko Sugioka, Takane Ito, Mitsuru Kawamura and Jun-ichi Shiota. 1999. Neurolinguistic evidence for rule-based nominal suffixation. *Language* 75. 739–63.

Hale, Kennth L. and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Ken Hale and Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvian Bromberger*. 53–109. Cambridge, MA: MIT Press.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and pieces of inflection. In Ken Hale and Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvian Bromberger*. 111–176. Cambridge, MA: MIT Press.

Harley, Heidi. 2008. On the causative construction. In Shigeru Miyagawa and Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 20–53 Oxford: Oxford University Press.

Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Iida, Masayo. 2002. Meishika-genshō no goi-ron-teki kōsatsu [A lexicalist study on nominalization phenomena] In Takane Ito (ed.), *Bunpō riron: Rekishikon to tōgo*. [Grammatical theory: The lexicon and syntax], 9–32. Tokyo: University of Tokyo Press.

Inoue, Kazuko. 1978. 'Tough sentences' in Japanese. In John Hinds and Irvin Howard (eds.), *Problems in Japanese syntax and semantics*, 122–154. Tokyo: Kaigakusha.

Ito, Takane. 2002. Nijū mekanizumu moderu to goijōhō no 'keishō' : Eigo no meishika no baai [Dual mechanism model and inheritance of lexical information: The case of English nominalization]. In Takane Ito (ed.), *Bunpō riron: Rekishikon to tōgo* [Grammatical theory: The lexicon and syntax], 225–248. Tokyo: University of Tokyo Press.

Ito, Takane and Yoko Sugioka. 2002. *Go no shikumi to gokeisei* [Mechanism of words and word formation]. Tokyo: Kenkyusha.

Ito, Takane, Yoko Sugioka and Hiroko Hagiwara. 2009. Neurological evidence differentiates two types of Japanese causatives. In Hiroto Hoshi (ed.), *The dynamics of the language faculty: Per spectives from linguistics and cognitive neuroscience*, 273–291. Tokyo: Kurosio Publishers.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.

Jacobsen, Wesley M. 1992. *The transitive structure of events in Japanese*. Tokyo: Kurosio Publishers.

Jacobsen, Wesley M. forthcoming. The semantic basis of Japanese transitive-intransitive derivational patterns. In Taro Kageyama and Wesley M. Jacobsen (eds.), *Transitivity and valency alterna tions: Studies on Japanese and beyond*. Berlin: De Gruyter Mouton.

Joanisse, Marc F. and Mark S. Seidenberg. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Science*s 96. 7592–7597.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–58.

Kageyama, Taro. 1989. The place of morphology in the grammar. In Geert Booij and Jan van Marle (eds.), *Yearbook of morphology* 2, 73–94. Dordrecht: Foris Publications.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron　Gengo to ninchi no setten* [Semantics of verbs: Interface of language and cognition]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1997. Denominal verbs and relative salience in lexical conceptual structure. In Taro Kageyama (ed.), *Verb semantics and syntactic structure*. 45–96. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2001. Word plus: The intersection of words and phrases. In Jeroen van de Weijer and Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*. 245–276. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Michiaki Saito, this volume, Chapter 1: Vocabulary strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kiparsky, Paul. 1997. Remarks on denominal verbs. In Alex Alsina, Joan Bresnan and Peter Sells (eds.), *Complex predicates*. 473–499. Stanford: CSLI Publications.

Kishimoto, Hideki. 2006. Japanese syntactic nominalization and VP-internal syntax. *Lingua* 116. 771–810.

Kitagawa Yoshihisa. 1986. More on bracketing paradoxes. *Linguistic Inquiry* 17. 177–183.

Kobayashi Hideki, Kiyo Yamashita and Taro Kageyama. this volume. Chapter 3: Sino-Japanese words. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: MIT dissertation. [Published by Garland, 1979.]

Kuroda, S.-Y. 1973. Where epistemology, style, and grammar meet: A case study from Japanese. In Stephen Anderson and Paul Kiparsky (eds.), *Festschrift for Morris Halle*. 377–391. New York: Holt, Rinehart and Winston.

Kuroda, S.-Y. 1987. Movement of noun phrase in Japanese. In S.-Y. Kuroda 1992. *Japanese syntax and semantics*, 253–292. Dordrecht: Kluwer Academic.

Kutas, Marta and Kara D. Federmeier. 2009. Event-related brain potential (ERP) studies of sentence processing. In M. Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*. 385–406. Oxford: Oxford University Press.

Lieber, Rochelle. 1992. *Deconstructing morphology*. Chicago: University of Chicago Press.

Lieber, Rochelle. 1998. The suffix *ize* in English: Implications for morphology. In Steven G. Lapointe, Diane K. Brentari and Patrick M. Farrell (eds.), *Morphology and its relation to phonology and syntax*. 12–33. Stanford: CLSI Publications.

Manning, Christopher D., Ivan A. Sag and Masayo Iida. 1999. The lexical integrity of Japanese causatives. In Robert D. Levine and Georgia M. Green (eds.), *Studies in contemporary phrase struc ture grammar*, 39–79. Cambridge: Cambridge University Press.

Marantz, Alec. 2007. Phases and words. In Sook-Hee Choe (ed.), *Phases in the theory of grammar*, 191–222. Seoul: Dong-In Publishing Co.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Matsumoto, Yo. forthcoming. Phonological and semantic subregularities in inchoative/noncausative-causative verb pairs in Japanese. In Taro Kageyama and Wesley M. Jacobsen (eds.), *Transitivity and valency alternations: Studies on Japanese and beyond*. Berlin: De Gruyter Mouton.

Miyagawa, Shigeru. 1989. *Syntax and Semantics 22: Structure and case marking in Japanese*. San Diego: Academic Press.

Miyagawa, Shigeru. 1999. Causatives. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*. 236–268. Oxford: Blackwell.

Namiki, Takayasu and Taro Kageyama. this volume. Chapter 6: Word structure and headedness. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word forma tion*. Berlin: De Gruyter Mouton.

Narrog, Heiko. forthcoming. Japanese transitivity pairs through time: A historical and typological perspective. In Taro Kageyama and Wesley M. Jacobsen (eds.), *Transitivity and valency alterna tions: Studies on Japanese and beyond*. Berlin: De Gruyter Mouton.

Nishio, Toraya. 1961. Dōshi no ren'yō-kei ni kansuru ichi-kōsatsu [A note on adverbal form of verbs.] *Kokugogaku* 43. 60–81. [Reprinted in Saito Michiaki and Masahiko Ishi (eds.), 1997 *Gokōsei* [Word structure], 192–212. Tokyo: Hituzi Syobo].

Ono, Naoyuki. this volume. Chapter 17: Agent nominals. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.

Pinker, Steven and Michael Ullman. 2002. The past and future of the past tense, *Trends in Cognitive Sciences* 6. 456–463.

Plag, Ingo. 1999. *Morphological productivity*. Berlin: Mouton de Gruyter.

Pollard, Carl and Ivan A. Sag. 1994. *Head Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Roeper, Thomas. 2005. Chomsky's *Remarks* and the transformationalist hypothesis. In Rochelle Lieber and Pavol Štekauer (eds.), *Handbook of word formation*, 125–146. Dordrecht: Kluwer.

Sadock, Jerrold M. 2012. *The modular architecture of grammar*. Cambridge: Cambridge University Press.

Sakakura, Atsuyoshi. 1966. *Gokōsei no kenkyū* [A study on word structure]. Tokyo: Kadokawa Shoten.

Saito, Michiaki and Masahiko Ishii (eds.). 1997. *Gokōsei* [Word structure]. Tokyo: Hituzi Syobo.

Selkirk, Elizabeth. 1982. *Syntax of words*. Cambridge, MA: MIT Press.

Shibatani, Masayoshi. 1976. Causativization. In Masayoshi Shibatani (ed.), *Syntax and Semantics 5: Japanese Generative Grammar*, 239–294. New York: Academic Press.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar. *Language* 64. 451–484.

Siegel, Dorothy. 1974. *Topics in English morphology*. Cambridge, MA: MIT dissertation.

Suga Kazuyoshi and Emiko Hayatsu (eds.). 1995. *Dōshi no jita* [Transitivity of verbs]. Tokyo: Hituzi Syobo.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English.* (The University of Chicago dissertation, 1984) New York: Garland.

Sugioka, Yoko. 1992. On the role of argument structure in nominalization. *Language, Communica tion and Culture* 10. 53–80. Yokohama: Keio University.

Sugioka, Yoko. 2002. Keiyōshi kara hasei suru dōshi no jita o megutte [On the transitivity alterna-tion of deadjectival verbs]. In Takane Ito (ed.), *Bunpō riron: Rekishikon to tōgo* [Grammatical theory: The lexicon and syntax], 91–116. Tokyo: University of Tokyo Press.

Sugioka, Yoko. 2011. Nominalization affixes and the multi-modular nature of word formation. In Etsuyo Yuasa, Tista Bagchi and Katharine Beals (eds.), *Pragmatics and autolexical grammar: In honor of Jerry Sadock*, 143–162. Amsterdam: John Benjamins.

Sugioka, Yoko, Takane Ito and Hiroko Hagiwara. 2001. Computation vs. memory in Japanese causative formation: Evidence from agrammatic aphasics. *Cognitive Studies* 8(1). 37–62.

Tagawa, Takumi. 2008. *Bunsan keitairon niyoru dōshi no katsuyō to gokeisei no kenkyū* [Study on verb inflection and word formation in Distributed Morphology.] Tsukuba: University of Tsukuba dissertation.

Teramura, Hideo. 1984. *Nihongo no shintakusu to imi II* [Syntax and semantics of Japanese II]. Tokyo: Kurosio Publishers.

Ullman, Michael T. 2001. A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience* 2. 717–726.

Vance, Timothy, J. 1987. *An introduction to Japanese phonology.* Albany: State University of New York Press.

van der Lely, Heather, K. J. and Steven Pinker. 2014. The biological basis of language: Insight from developmental grammatical impairments. *Trends in Cognitive Sciences* 18. 586–595.

Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* 12. 245–274.

Yamazaki, Kaoru. 1992. *Keiyōshi jodōshi no kenkyū* [A study on adjectives and auxiliary verbs]. Osaka: Izumi Shoin.

Yumoto, Yoko. this volume, Chapter 9: Conversion and deverbal compound noun. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kentaro Nakatani

11 Complex predicates with *-te* gerundive verbs

1 Introduction

Japanese abounds with expressions comprising multiple verbs, as evident in the rich inventory of morphological compounds consisting of two verbs (see Chapter 8 [Kageyama, this volume] for review). Another class of verbal complexes is the complex predicates composed of the first verb with the "gerundive" *-te* suffix, followed by a lexically designated set of "auxiliary-like" verbs. In (1), for example, the gerundive verb *yat-te* (< *yar-* 'do' + *te*) shows up in the first (V$_1$) position and the "auxiliary-like" verbs *oku* 'put' and *iru* 'be' in the second (V$_2$) position.

(1) a. *Watasi wa syukudai o* **yat-te oi-ta**.
      I   TOP homework ACC do-GER put-PST
      'I have done the homework.'

   b. *Taroo wa geemu o* **yat-te i-ru**.
      Taro TOP game ACC do-GER exist-PST
      'Taro is playing a game.'

In (1a), *yat-te oita* as a whole represents the purposeful completion of an activity with the meaning of 'have done (for some later purpose)', and in (1b) *yat-te iru* denotes an ongoing activity with the meaning of 'is doing'. In this chapter, we will call this type of verbal complex the "V$_1$-*te* V$_2$ complex predicate" or simply, the "V$_1$-*te* V$_2$ predicate".

    Note that the gerund verb ending in *-te*, which putatively originates from some kind of perfective inflection or a conjunctive particle, has a variety of usages in the coordination and subordination of clauses, as in *wain o non-de, neru* [wine ACC drink-GER sleep] 'to drink wine and go to bed' and *kawa o oyoi-de wataru* [river ACC swim-GER cross] 'to cross the river, swimming'. Unlike these coordinating and subordinating functions, which are not specifically constrained by the lexical selection of verbs, the range of V$_1$-*te* V$_2$ predicates under discussion are uniquely delimited by a class of lexically specified verbs that are qualified for the V$_2$ position.

    Apart from the presence of the gerundive ending, V-*te* V predicates have a number of properties that suggest that they are distinct from both lexical and syntactic V-V compound verbs. For example, only a limited number of verbs (approximately twelve, as exemplified in Table 1 below) can enter the V$_2$ position in V$_1$-*te* V$_2$ complexes, whereas a much greater variety is attested in the lexical VV compounds;

in this respect, V-*te* V complexes are similar to syntactic V-V compounds, in which the verbs that enter in the V_2 position are limited to about 30 verbs. Like the syntactic V-V compounds, the choice of V_1 in V_1-*te* V_2 predicates is generally unrestricted, a property not found in the lexical VV compounds. The fact that idioms and syntactically complex units are allowed in V_1 in V_1-*te* V_2 complexes suggests that the sequence of V_1-*te* V_2 does not constitute a lexical unit (more on this in Section 3.1). Furthermore, in V_1-*te* V_2 predicates, only V_2 undergoes semantic lightening,[1] whereas the effects of semantic lightening are diverse in V_1 and V_2 in the lexical V-V compounds. In what follows, the verbs that appear in the V_2 position of V-*te* V complex predicates will be referred to as "aux V_2" hereafter, following the way they are labeled in the literature (*hojo dōshi* 'helping verb': Matsushita 1930; Yoshikawa 1973; Martin 1975; Takahashi 1976; Morita 1977; Masuoka 1992; Miyake 2005, etc.).

The twelve verbs that occur as V_2 are listed in Table 1 together with their original and auxiliary meanings (see Section 2 for example sentences).

Table 1: List of twelve verbs that can enter the V_2 position in the V_1-*te* V_2 predicate

V_2 verbs	Original meanings	Approximate meanings as aux V_2s
kuru	'to come'	Coming/going motion, a gradual change/
iku	'to go'	advancement of state/event, etc.
iru	'(for an animate) to exist'	Progressive or perfect aspect
aru	'(for an inanimate) to exist'	Perfect aspect, with a connotation of preparation
miru	'to see'	Attempt
miseru	'to show'	Demonstration
kureru/ageru	'to give (to the speaker)' 'to give (to a non-speaker)'	Benefactive activity
morau	'to receive/be given'	Receipt of a benefit
oku	'to put'	Preparation
simau	'to put away'	Emphasis on completion
hosii (adjective)	'to want'	Wish for someone to do something

Of special importance in this table is that the original meanings are more or less retained in the auxiliary semantics, thus strongly suggesting lexical relatedness between the auxiliary usage and the corresponding usage as a lexical verb. This relatedness cannot be captured by simply treating the twelve verbs as "auxiliaries".

---

1 An exception to this generalization is *yat te kuru* [do GER come], in which *yat te* is semantically almost (if not completely) vacuous, making the whole expression approximately synonymous to simple *kuru*, although the former invokes a slight connotation of coming 'all the way'. Interestingly, the *iku* counterpart, *yat te iku* [do GER go] does not have this type of motional interpretation, and only means 'to get along'.

In fact, unlike the English modal auxiliaries, the Japanese aux V_2s behave morphologically in exactly the same way as ordinary lexical verbs in exhibiting tense and other inflections. This fact provides good motivation for exploring a derivational approach where the auxiliary status of V_2 is assumed to be synchronically associated with its main verb counterpart. Another reason a derivational approach is taken is that we aim not only to simply describe the various properties of the V-*te* V constructions, but also to construct a theory that would make predictions based on the semantics of the parts (i.e. V_1, *-te*, and V_2). In Sections 3 and 4, we examine the morphosyntactic and semantic properties of the V-*te* V predicate that cannot be properly accounted for by simply assuming that aux V_2s are fixed as auxiliaries independent of their main-verb counterparts. Rather, it is necessary to take into account the interaction between the semantics of V_1-*te* (more precisely, VP_1-*te*) and the original, main-verb semantics of V_2 to capture the whole picture of the phenomena.

This chapter is organized as follows. In Section 2, we overview various instances of the V-*te* V predicate, sorting them by the valency properties of V_2s: intransitive verbs (*kuru* 'to come', *iku* 'to go', *iru* '(for animates) to exist', *aru* '(for inanimates) to exist'), transitive and ditransitive verbs (*kureru* 'to give (to the speaker)', *ageru* 'to give (to a non-speaker)', *morau* 'to be given/to receive'), *oku* 'to put', *simau* 'to put away', *miru* 'to see', *miseru* 'to show'), and an adjective (*hosii* 'to want').[2] In Section 3, we analyze the morphosyntactic properties of the complexes, and in Section 4, their semantic issues are discussed. Section 5 summarizes the chapter.

2 The V-*te* V constructions: Overview

This section provides examples of the V-*te* V predicate, reviewing the semantics of the aux V_2s in comparison with that of their main-verb counterparts. The V_2s may be either intransitive or transitive/ditransitive. There is also an adjective, *hosii* 'to want', that occurs with a gerundive verb to form a complex predicate.

2.1 Intransitive V_2s

We will first look into intransitive aux V_2s, starting with the motion verbs *kuru* 'to come' and *iku* 'to go', shown in (2) along with their main-verb usage:

---

2 Some other verbs such as *kaeru* 'to go back', *mawaru* 'to go around', and *yokosu* 'send' can argu ably enter the V_2 position in the V_1 *te* V_2 complex with slightly less productivity. See Nakatani (2013: 22 24) for further details.

(2) a. *Taroo ga ofisu ni* { *ki-ta* / *it-ta* }.
    Taro NOM office DAT { come-PST / go-PST}
    'Taro {came/went} to the office.'

  b. *Taroo ga hon o kat-te* { *ki-ta* / *it-ta* }.
    Taro NOM book ACC buy-GER { come-PST / go-PST }
    'Taro bought a book and came/went (away); Taro bought a book and
    brought/took it along.'

The verbs of coming and going are used as main verbs in (2a) while they are used as aux V_2s as part of a V-*te* V complex in (2b), although their auxiliary-like status may not be immediately obvious from the glosses and translations. One reason for this obscurity is that the sense of directional movement is not lost in the auxiliary usage in (2b). However, there are reasons to believe that the V_1-*te* V_2 forms a complex predicate, as discussed in Section 3.3.

    These directional motion verbs can also be used in a non-physical sense in the V-*te* V frame, in which case they are often labeled as aspectual markers in the literature (Teramura 1984; Yoshikawa 1973; Shibatani 2007; among many others):

(3) a. *Zutuu ga si-te* { *ki-ta* / **it-ta* }.
    headache NOM do-GER { come-PST / *go-PST }
    'I am getting a headache.'

  b. *Syoonen wa dandan tuyoku nat-te* { *ki-ta* / *it-ta* }.
    boy TOP gradually strong become-GER { come-PST / go-PST }
    'The boy was getting stronger and stronger.'

These examples do not depict physical movements of the subject entities; rather, the verbs of coming/going have shifted to aspectual meanings designating the gradual unfolding of the events denoted by the verbs in V_1, in much the same way that English *come* and *go* take on aspectual meaning when followed by adjectives, as in *come loose* and *go sour*. For this reason, *kuru* and *iku* in (3) are commonly regarded as aspectual markers in the literature of Japanese grammar (Takahashi 1976; Teramura 1984). The directionality contrast between *kuru* and *iku* is still retained in these non-physical examples. For example, (3a) depicts that the speaker is getting a headache, and the sense of getting a headache is most naturally described as "coming" to the speaker; for this reason, it would be very awkward to replace *kuru* with *iku* in this case. By contrast, from the speaker's perspective, the gradual change in "the boy" depicted in (3b) can be naturally construed as either "coming" (the speaker comes to realize the change in the boy's condition more and more clearly) or "going" (the speaker perceives the boy's change as a progress toward a better future).

In the above cases, the V$_2$s introduce a sense of on-going, gradual changes, so they can safely be regarded as aspectual markers. However, not all non-physical *kuru* 'come' can be regarded as an aspectual marker. Let us consider the following pair.

(4) a. *Taroo ga   boku ni   tegami o   okut-te   ki-ta.*
       Taroo NOM I   DAT letter   ACC send-GER come-PST
       'Taro sent me a letter.'

   b.?*Taroo ga   boku ni   tegami o   okut-ta.*
       Taroo NOM I   DAT letter   ACC send-PST
       'Taro sent me a letter.'

Unless a comma pause is put after *okut-te* [send-GER], sentence (4a) cannot mean that Taro himself literally came to the speaker location, indicating that *kuru* 'come' in this example does not denote a physical motion. However, the verb *kuru* in (4a) is not purely aspectual, either, because it does not designates a gradual unfolding of the event of Taro's sending me a letter. In fact, the aspectual meaning of (4a) does not change if the verb of coming is removed, as in (4b) (which is a little awkward but acceptable). From this, it is concluded that *kuru* in (4a) is neither physical nor aspectual.[3]

Two other intransitive verbs that may enter the V$_2$ position of complex predicates are verbs of existence, *iru* and *aru*, which are purely stative verbs. When used as main verbs, they contrast in terms of the animacy of their subject entities, the former being restricted to animate subjects and the latter to inanimate ones.

(5) a. *Soko   ni   neko ga   { i-ru   / *ar-u }.*
       there DAT cat   NOM { exist PRS / *exist PRS }
       'There is a cat over there.'

   b. *Soko   ni   hon   ga   { *i-ru   / ar-u }.*
       there DAT book NOM { *exist PRS / exist PRS }
       'There is a book over there.'

Although the two verbs in their usage as lexical verbs are very similar in terms of semantics except for the animacy restriction, they may be semantically contrastive when used as aux V$_2$s, where the animacy contrast disappears. Observe the example in (6).

---

3 Interestingly, *iku* cannot successfully replace *kuru* in (4a). See Imani (1990); Nakatani (2008) for discussions on this matter.

(6) *Taroo wa syukudai o yat-te* { *i-ru* / *ar-u* }.
    Taro TOP homework ACC do-GER { exist PRS / exist PRS }
    'Taro { is doing the homework / has done the homework }.'

The distinction of *iru* and *aru* is reflected in the difference in the aspectual inter-pretation: the prominent reading of V-*te iru* in this example is progressive, while V-*te aru* represents a perfective aspect (see also Chapter 15 [Jacobsen, this volume] for further discussions on aspect).

What makes the matter complicated is that, besides the progressing meaning, *V-te iru* also allows a perfective interpretation, especially when there is support from the context or a perfective adverbial, in which case its semantics is truth-conditionally indistinguishable from the *aru* version. In (7), *iru* and *aru* are rendered into English as "Taro has already done his homework'. In Section 4.3, it will be shown that the two verbs differ in that *aru*, but not *iru*, has an additional connota-tion like 'ready for future use'.

(7) *Taroo wa syukudai o sudeni yat-te* { *i-ru* / *ar-u* }
    Taro TOP homework ACC already do-GER { exist PRS / exist PRS }
    'Taro has already done the homework.'

Finally, we note that the V$_1$-*te ar-u* construction has a prominent property of valency change in syntactic structure – a property that is usually associated with passive constructions.

(8) *Syukudai ga yat-te* { *ar-u* / **i-ru* }.
    homework NOM do-GER { exist PRS / *exist PRS }
    'The homework has already been done.'

In (8), where only *aru* is acceptable, the agent ('Taro') disappears and the object ('homework') is promoted to the subject of the whole V$_1$-*te ar-u* complex. We return to this issue in Section 4.3.

2.2 Transitive V$_2$s

Now we turn to the transitive V$_2$s, whose lexical counterparts are, with the exception of *miru* 'to see', are all ditransitive or three-place predicates: *kureru* 'to give', *ageru* (*yaru* in casual speech)[4] 'to give', *morau* 'to be given', *oku* 'to put', *simau* 'to put

---

4 *Yaru* in Modern Japanese (which is homophonous with *yaru* 'to do') is a variant of *ageru*, and is only used when the agent gives something to a socially lower ranked person than the agent, or to nonhumans. By contrast, *ageru* (derivative of *ageru* in the sense of "raise") is a polite version of it, which has a wider usage than *yaru*, especially with younger speakers. In most cases, *ageru* can

away', and *miseru* 'to show'. When used as aux V$_2$s in the V$_1$-*te* V$_2$ frame, these verbs never take their own object independent of the object argument of V$_1$, and instead, the event or state associated with VP$_1$ is interpreted as the "hidden" direct object in many cases. Therefore, on the surface, it appears as if V$_2$ were taking VP$_1$ as its complement. However, there are reasons to doubt such a view, as we shall briefly touch on later in this section and discuss in more detail in Section 4.3. Let us first overview transitive/ditransitive examples, starting with the verbs of giving.

Japanese has two lexical verbs of giving, *kureru* and *ageru* (or colloquial *yaru*), which are contrastive with respect to directionality relative to the speaker. That is, in parallel with the contrast between deictic motion verbs *kuru* 'come' and *iku* 'go', *kureru* 'to give to the speaker' represents motion toward the speaker, and *ageru* (or *yaru*) motion away from the speaker.

(9) a. *Taroo ga boku ni hon o* { *kure-ta* / *age-ta }.
    Taro NOM me DAT book ACC { give-PST / *give-PST }
    'Taro gave a book to me.'

   b. *Boku ga Taroo ni hon o* { *kure-ta / *age-ta* }.
    I NOM Taro DAT book ACC { *give-PST / give-PST }
    'I gave a book to Taro.'

Furthermore, Japanese has another verb of giving, *morau*, which designates the same event as described by *kureru* (9a) from an inverse perspective (i.e. from the perspective of the speaker or a person closely associated with him). I tentatively gloss this verb as 'to be given' even though it is an active verb, partly because its semantics is closely related to the notion of possession transfer (and thus distinct from similar verbs such as *uketoru* 'receive' and *eru* 'to obtain') and partly because it has some syntactic and semantic properties similar to the passive (for example, the agent of the event can be dative-marked). Furthermore, the two verbs of giving, *ageru* and *kureru*, cannot be passivized, possibly as a result of lexical blocking by *morau*. An example of main-verb *morau* is shown in (10).

(10) *Taroo ga Hanako ni hon o morat-ta.*
    Taro NOM Hanako DAT book ACC be.given-PST
    'Taro obtained a book from Hanako.'

---

replace *yaru*, but not vice versa. However, it should be noted that the V *te yaru* construction has not only a benefactive use but also an anti benefactive, adversarial use, which is not attested in the V *te ageru* construction. For example: *Korosi te yar u* '(lit.) kill GER give PRS = 'I'll kill you' and *Zettai gookaku si te yar u* '(lit.) certainly pass do GER give' = 'I'll definitely pass the exam (and prove that you were wrong!)'. See Komatsu (1964) and more recently, Yamada (2004) for further discussion.

All these three verbs can enter the V-*te* V frame, as illustrated below:

(11) a. *Taroo ga   hon   o   kat-te   { kure-ta / age-ta (yat-ta) }.*
     Taro   NOM book ACC buy-GER { give-PST / give-PST }
     'Taro bought a book (for the speaker / for a non-speaker).'

  b. *Taroo ga   haha   ni   hon   o   kat-te   morat-ta.*
     Taro   NOM mother DAT book ACC buy-GER be.given-PST
     'Taro benefited from his mother having bought a book.'
     (Taro's mother bought him a book.)

In the above examples, a sense of physical theme is retained for the verbs of giving (i.e. the book was given). However, more abstract benefit can be transferred as well.

(12) a. *Taroo ga   paatii ni   ki-te   { kure-ta / age-ta (yat-ta) }.*
     Taro   NOM party DAT come-GER { give-PST / give-PST }
     'Taro came to the party (benefitting the speaker / a non-speaker).'

  b. *Watasi wa   Taroo ni   paatii ni   ki-te   morat-ta.*
     I   TOP Taro   DAT party DAT come-GER be.given-PST
     'I had Taro come to the party. (I benefited from Taro's coming to the party)'

(12) does not imply transfer of any physical object but only the receipt of an abstract favor or benefit resulting from the giving event.

   Two other ditransitive verbs of causative theme movement, albeit without the sense of possession transfer, can participate in V-*te* V complexes. They are *oku* 'to put' and *simau* 'to put away', which denote the action of placing some physical object onto a two-dimensional surface and into a three-dimensional storage space, respectively.

(13) a. *Taroo wa   kaban o   yuka ni   { oi-ta   / *simat-ta }.*
     Taro   TOP bag   ACC floor DAT { put-PST / *put.away-PST }
     'Taro { put / *put away } the bag on the floor.'

  b. *Taroo wa   kaban o   kinko ni   { ??oi-ta   / simat-ta }.*
     Taro   TOP bag   ACC safe   DAT { ??put-PST / put.away-PST }
     'Taro { ??put / put away } the bag into the safe.'

As for (13a), it is impossible to conceive of a floor as a storage space, leading to the unacceptability of *simau* 'to put away'. On the other hand, *kinko* 'a safe' in (13b) is a typical storage device, for which *simau* is more appropriate. Locations such as *tana* 'shelf' readily allow both storage and surface interpretations, and are thus compatible with either of the two verbs.

When used in the V-*te* V frame, both *oku* 'put' and *simau* 'to put away' lose their physical movement senses, although the contrastive sense of location is retained:

(14) *Taroo wa syukudai o yat-te* { *oi-ta* / *simat-ta* }.
    Taro TOP homework ACC do-GER { put-PST / put.away PST }
    'Taro did the homework.'

In (14), both *yat-te oita* and *yat-te simatta* denote the same event, with the difference lying in the connotation. With *oku* (past tense *oi-ta*), it is conventionally implicated that Taro has done the homework, probably well in advance, so that he is ready for the next move (e.g. he can submit the homework whenever requested, he can work on other things, etc.). Because of this "readiness" connotation, *oku* as an aux V_2 is sometimes called an auxiliary verb of "attitude" or "planning" (Takahashi 1976; among others). With *simau* (past tense *simat-ta*), on the other hand, it is felt that the completion of the VP_1 event is emphasized.

    Because of this intuition, *simau* in the V-*te* V frame is often called a completive aspectual marker in the literature of traditional Japanese grammar (Takahashi 1976; Teramura 1984; among many others). Some researchers such as Teramura (1984: 123) further state that *oku* is not an aspectual marker in the same sense that *simau* is. However, such an impressionistic distinction is not well founded. In the above example, even V-*te oku* designates the completion of VP_1 (note that "readiness" presupposes completion). Furthermore, omission of *oku/simau* in (14) would not lead to a change in the temporal/aspectual interpretation of the sentence, as illustrated in (15), where no aux V_2 is involved.

(15) *Taroo wa syukudai o yat-ta*.
    Taro TOP homework ACC do-PST
    'Taro did the homework.'

The illusion that *oku* and *simau* function differently in terms of aspectual interpretation possibly stems from the fact that the former invokes a sense of readiness, which is more psychological than aspectual, while the latter does not. However, it is known that not only *oku* but also *simau* may conventionally implicate some psychological state of the speaker. For example, the sense of completion may be accompanied by the sense of regret when the completion of VP_1 is undesirable for the speaker.

(16) *Taroo wa kabin o wat-te* { #*oi-ta* / *simat-ta* }.
    Taro TOP vase ACC break-GER { #put-PST / put.away PST }
    'Taro broke the vase.'

In (16), the use of *simau* invokes an interpretation where the speaker must be considering the event as very unfortunate and not undoable. The use of *oku* (past *oi-ta*)

in place of *simau* (past *simat-ta*) sounds awkward here because it is pragmatically difficult (though not impossible) to imagine a situation in which the completion of Taro's breaking the vase is intended as a premeditated preparation for the next move. All in all, there seems to be no strong evidence for distinguishing the two verbs in terms of their status as an "aspectual marker", which neither is.

Finally, let us examine the verb of visual perception, *miru* 'to see, look at, or watch', and its causative version, *miseru* 'to show'.

(17) *Taroo wa bideo o*   { *mi-ta*   / *mise-ta* }
　　 Taro   TOP video ACC { see-PST / show-PST }
　　 'Taro { saw / showed } the video.'

These verbs can enter the V-*te* V frame without the visual sense of seeing:[5]

(18) *Taroo wa uokka o   non-de*   { *mi-ta*   / *mise-ta* }.
　　 Taro   TOP vodka ACC drink-GER { see-PST / show-PST }
　　 'Taro tried drinking vodka / Taro demonstrated that he could drink vodka.'

As can be seen in the above example, V_1-*te miru* 'to see' roughly means 'try V_1-ing' and V_1-*te miseru* 'show off by V_1-ing'. Note that the former does *not* mean 'try to V_1'; rather, it means 'to actually do V_1 and see what its result would be like'. Thus, as a paraphrase of the *mita* version of (18), (20) is more appropriate than (19):

(19) *Taroo wa* [*zibun ga   uokka o   nom-u   tokoro   o*]   *mi-ta*.
　　 Taro   TOP [self   NOM vodka ACC drink-PRS occasion ACC] see-PST
　　 'Taro saw himself drinking vodka.'

(20) *Taroo wa* [*zibun ga   uokka o   non-da   kekka o*]   *mi-ta*.
　　 Taro   TOP [self   NOM vodka ACC drink-PST result ACC] see-PST
　　 'Taro saw the result of himself drinking vodka.'

This shows that despite its appearance, (18) does not involve simple complementation because the vodka-drinking event itself is not the complement of *mita*. We shall return to the issue of what I might call "complementation illusion" in Section 4.3.

---

5 Gerundive particle *te* is phonetically realized as *de* when preceded by either (i) a verb root ending with a nasal (e.g. *yom* 'to read' + *te* = *yon de*), (ii) a verb root ending with a voiced bilabial stop /b/ (e.g. *yob* 'to call' + *te* = *yon de*, homophonous with the gerundive form of *yom*), or (iii) a verb root ending with a voiced velar stop (e.g. *kag* 'to sniff' + *te* = *kai de*). Note that the same rules apply to past tense marker *ta*.

2.3 Desiderative adjective with a gerundive V₁

In Japanese, three types of desiderative notions are linguistically distinguished: (i) one wants something; (ii) one wants to do something; and (iii) one wants somebody to do something. For the first case, the adjective *hosii* 'want, be in need of' is used, and for the second case, the bound suffix *-tai* 'be eager to' is used:

(21) *Watasi wa hon ga hosi-i.*
　　 I　　　TOP book NOM want-PRS
　　 'I want a book.'

(22) *Watasi wa hon { ga / o }　 yomi-ta-i.*
　　 I　　　TOP book { NOM / ACC } read-DESI PRS
　　 'I want to read a book.'

For the third case, *hosii* is combined with V₁-*te*, and the embedded subject is marked with the dative particle *ni*.

(23) *Watasi wa Hanako ni　 hon　 o　　yon-de　 hosi-i*
　　 I　　　TOP Hanako DAT book ACC read-GER want-PRS
　　 'I want Hanako to read a book.'

One interesting fact is that *hosii* itself is not capable of assigning dative case; a dative NP may appear only when it enters the gerundive construction.

(24) **Watasi　wa　Hanako　ni　　hosi-i.*
　　　I　　　TOP　Hanako　DAT　want-PRS

Another interesting property of the V-*te hosii* construction is that the embedded subject may be nominative-marked in some cases.

(25) *Watasi wa yuki { ga　 / ??ni }　 hut-te　 hosi-i.*
　　 I　　　TOP snow { NOM / ??DAT} fall-GER want-PRS
　　 'I hope that it snows.' (lit. 'I want snow to fall.')

The degraded acceptability of the dative *ni* in the above seems to be associated with the fact that the embedded subject is inanimate. In contrast, (23), which has an animate embedded subject, would sound very odd if the embedded subject is nominative-marked. However, it is not the case that an animate embedded subject is always incompatible with nominative-marking. When a "candidacy" interpretation (or "exhaustive listing" in Kuno's (1973) terminology) is readily available, the embedded animate subject may be nominative.

(26) *Watasi wa Koizumi* { *ga* / *ni* } *syusyoo ni nat-te*
     I   TOP Koizumi { NOM / DAT } prime.minister DAT become-GER

     hosi-i.
     want-PRS
     'I want Koizumi to be the prime minister.'

Matsumoto (1996), Muraki (1978), and Nakatani (2013: 153–156) discuss this issue in greater detail.

2.4 Summary

In this section, we have quickly overviewed the major instances of the V$_1$-*te* V$_2$ predicates, and observed the auxiliary-like semantics of V$_2$s. Some of these V$_2$s are originally intransitive, and others are transitive or ditransitive. None of them seem to take an accusative object independent of V$_1$ when used in the V$_1$-*te* V$_2$ frame, leading to an illusion that the transitive V$_2$s take VP$_1$-*te* as a complement (more on this in the next section).

Most V$_2$s with the meaning of physical movement seem to lose their physical semantics when used as aux V$_2$s, although *kuru* 'to come' and *iku* 'to go' show a wide variety of meaning including physical movement even as aux V$_2$s. The non-physical semantics of aux V$_2$s can be considered aspectual in some cases (*kuru* 'to come', *iku* 'to go', and *iru/aru* 'to exist'), although the majority, including V-*te oku* 'to put' and *simau* 'to put away', are not strictly aspectual (contrary to the popular belief in the literature). These properties are summarized in Table 2.

Table 2: Summary of the types of the V-*te* V constructions

Original valency	V-*te* V instances	Physical movement	Non-physical movement	Aspectual function
Intransitive	V-*te ku ru* 'to come' V-*te ik u* 'to go'	Yes	Yes	Yes
	V-*te i ru* 'to exist' V-*te ar u* 'to exist'	–	–	Yes
Transitive	V-*te mi ru* 'to see'	–	–	No
Ditransitive	V-*te mise ru* 'to show'	–	–	No
	V-*te kure ru* 'to give' V-*te age ru* 'to give'	(Yes)*	Yes	No
	V-*te morau* 'to be given'	(Yes)*	Yes	No
	V-*te ok u* 'to put' V-*te sima u* 'to put away'	No	Yes	No
Adjective	V-*te hosi i*	–	–	(Yes)**

*The physical movement reading may be merely an inference from the non-physical, benefactive reading. See Section 4.3.
**"Yes" only in the sense that desideratives introduce intensional context.

3 The morphosyntax of V-*te* V predicates

In this section, we consider the morphological and syntactic properties of V-*te* V complex predicates, particularly focusing on the issue of the morphological "size" of the complex.

3.1 V-*te* V versus V-V

As briefly touched on in Section 1, V-*te* V predicates are distinct from lexical V-V compounds in several respects: (i) *-te* is present in the former; (ii) it is always V_2 that undergoes semantic lightening in the former whereas various semantic patterns are observed in the latter; (iii) the former allows idiomatic expressions and light-verb constructions in the V_1 position, which the latter does not. In these respects, the V-*te* V predicate is more similar to syntactic V-V compounds (Kageyama 1993). The set of examples in (27) illustrates the third point, that lexical V-V compounds do not allow a light verb construction such as *kikku-suru* 'kick$_N$-do' in the V_1 position (in contrast with simple lexical verb *keru* 'to kick') while the syntactic VV and V-*te* V constructions do.

(27) a. *Hanako ga   Taroo o   { ket-ta   / kikku-si-ta }.*
     Hanako NOM Taro   ACC { kick$_V$-PST / kick$_N$-do-PST }
     'Hanako kicked Taro.'

   b. Lexical VV:
     Hanako ga   Taroo o   { keri   / *kikku-si } korosi-ta.
     Hanako NOM Taro   ACC { kick$_V$ / *kick$_N$-do } kill-PST
     'Hanako kicked Taro to death.'

   c. Syntactic VV:
     Hanako ga   Taroo o   { keri   / kikku-si } hazime-ta.
     Hanako NOM Taro   ACC { kick$_V$ / kick$_N$-do } start-PST
     'Hanako started to kick Taro.'

   d. V-*te* V:
     Hanako ga   Taroo o   { ket-te   / kikku-si-te }   simat-ta.
     Hanako NOM Taro   ACC { kick$_V$-GER / kick$_N$-do-GER } put.away-PST
     'Hanako (regretfully) kicked Taro.'

(27d) illustrates only one instance of aux V_2 *simau* 'to put away', but any other aux V_2 could also accommodate complex forms such as *kikku-suru* in the V_1 position. Furthermore, unlike lexical V-V compounds, both syntactic -V and V-*te* V constructions readily allow V_1 to be passivized or causativized:

(28)   a.   Lexical V-V:
             *Ken ga   ker-are   korosi-ta.
             Ken NOM kick-PASS kill-PST

       (cf. Ken ga   keri   koros-are-ta.
            Ken NOM kick   kill-PASS PST
            'Ken was kicked to death.')

       b.   Syntactic V-V:
            Ken ga   ker-are   hazime-ta.
            Ken NOM kick-PASS start-PST
            'Ken started to be kicked.'

       c.   V-*te* V:
            Ken ga   ker-are-te   simat-ta.
            Ken NOM kick-PASS GER put.away-PST
            'Ken was (regretfully) kicked.'

It is obvious from the above observations that the V-*te* V predicate does not constitute a morphological "word": Rather, it is a syntactic unit.

In fact, there are some pieces of evidence that the V-*te* V complex is a syntactic unit even larger than syntactic V-V compounds. For example, Kageyama (1993: 360) observes that the V-*te* V predicate cannot enter the honorific template *o*-V-*ni nar-u*, as shown in (29), while syntactic V-V compounds can, as in (30):

(29) a. *Yamada-sensei   ga   sore o   **tabe-te**   **mi**-ta.*
        Yamada-teacher NOM it   ACC eat-GER   see-PST
        'Professor Yamada tried eating it.'

     b. [Honorific with the *o*-V-*ni nar-u* template]
        Yamada-sensei   ga   sore o   o*-[tabe-te mi**]-*ni nat-ta*

(30) a. *Yamada-sensei   ga   sore o   **tabe-hazime**-ta.*
        Yamada-teacher NOM it   ACC eat-begin-PST
        'Professor Yamada began eating it.'

     b. [Honorific with the *o*-V-*ni nar-u* template]
        Yamada-sensei   ga   sore o   o-[**tabe-hazime**]-*ni nat-ta.*

Based on the assumption that the honorific template *o*-V-*ni nar-u* is acceptable only when the V is a "word-like" unit, the above contrast shows that the V-*te* V complex is even less word-like than syntactic V-V compounds.

The possible attachment of various focus particles to the right of -*te* (Martin 1975: 510–512) is also a strong indicator that there is a phrasal boundary between

V1-*te* and V₂. This applies unexceptionally to all the instances of V-*te* V complexes, even to *teiru* (-*te iru*), which is a prominent aspectual marker in Japanese, as exemplified below.

(31) *Kanozyo wa nai-te* {*wa/mo/sae/nanka*}   *i-nakat-ta.*
       she   TOP cry-GER {PART/PART/PART/PART} exist-NEG PST
       'She was not {really/also/even/ever} crying.'

Since focus particles are inserted in syntax rather than in the lexicon, the acceptability of various focus particles shown in the curly brackets in (31) shows that the morphosyntactic boundary lies between *-te* and V₂.

3.2 Morphosyntax of *-te*

The status of the V-*te* V complex as a looser unit compared to the syntactic V-V compounds may be related to the existence of the "gerundive" particle *-te*, which is absent in the V-V compounds. In this subsection, the properties of *-te* in general (not just the one used in the V-*te* V predicate) are discussed.

Although *-te* is tentatively glossed as "gerundive" in this handbook, its syntactic distribution is wider than that of what is usually termed "gerundive" in European languages. It is best described as a suffixal clause linker that makes a VP or a clause an adjunct to another VP or a clause.

(32) a. *Haha   ga   [saihu o   mot-te]   nige-ta.*
           mother NOM [wallet ACC hold-GER] run.away-PST
           'The mother ran away holding her wallet.'

      b. [*Titi   ga   kaet-te*]   *musume wa yorokon-da.*
         [father NOM return GER] daughter TOP get.delighted-PST
         'The father came back, and the daughter got delighted.'

The bracketed phrase in (32a) specifies the manner of the main verb *nige-ta* 'ran away', and looks much like a gerund or converb. In addition to such a usage, *-te* can host a full clause with an overt subject, as shown in (32b), in which *-te* looks like a coordinator. However, this *-te* is obviously not a coordinator because the bracketed phrases in (32) can be scrambled:

(33) a. [*Saihu o   mot-te*]   *haha   ga   nige-ta.*
         [wallet ACC hold-GER] mother NOM run.away-PST
         'The mother ran away holding her wallet.'

b. *Musume wa* [*titi   ga   kaet-te*]   *yorokon-da.*
   daughter TOP [father NOM return GER] get.delighted-PST
   'The father came back, and the daughter got delighted.'

Therefore, it would be more appropriate to analyze *-te* as a clause-projecting verbal suffix, whose subject position can be null. Note that *-te* can be attached to a verb stem with passive and/or causative suffixes attached.

(34) *Hanako wa* [*mazui   mono o   tabe-sase-rare-te*]   *okot-ta.*
     Hanako TOP [unpalatable thing ACC eat-CAUS PASS GER] get.angry-PST
     'Hanako was made to eat an unpalatable thing and got angry.'

It is also notable that verb stems inflect with *-te* in the same way as they inflect with *-ta* (e.g. *kag(-u)* 'to sniff' + *-te* = *kai-de*; *kag(-u)* + *-ta* = *kai-da*), except that adjectives conjugate differently with *-te* and *-ta* (e.g. *kawai(-i)* 'cute' + *-te* = *kawaiku-te*; *kawai(-i)* + *-ta* = *kawaikat-ta*). Some researchers also point out semantic similarities between *-te* and *-ta* (Kuno 1973; Matsuo 1936; Nishida 1977; Yoshikawa 1973: but see also Hasegawa 1996). Finally, *-ta* marks a finite clause, whereas *-te* only appears in a non-finite environment (except for *-te* found in truncated constructions such as Imperative *Ki-te!* 'Please come!', which is a short form of *Ki-te kudasai!* where *kudasai* is a polite imperative form of *kure-ru* 'give (to the speaker)'). On these grounds, *-te* may be assumed to be a non-finite variant of *-ta*, which functions as a relative past tense marker, albeit defectively (Nakatani 2003a, 2013: 57–106). Under this analysis, *-te*, as well as *-ta*, is best analyzed as a tense or aspectual head, which is placed at the edge of a propositional projection. Let us assume, adopting the generativist assumption that tense projects a clause, that the bracketed phrases in (32)–(34) are projections of *-te* and call them *te* phrases, or *te*P.

 The fact that *te*P can be adjoined to various places makes it ambiguous whether a sequence of V$_1$-*te* and V$_2$ constitutes a complex predicate or a mere adjunction structure. For instance, *te*P in the following examples is placed in different positions without changing the propositional interpretation of the whole sentence:

(35) a. [$_{teP}$ *Kaban o   katui-de*]   *Taroo wa tyuusyazyoo ni   ki-ta.*
       [$_{teP}$ bag   ACC shoulder-GER] Taro TOP parking.lot DAT come-PST
       'Taro came to the parking lot, shouldering a bag.'

   b. *Taroo wa* [$_{teP}$ *kaban o   katui-de*]   *tyuusyazyoo ni   ki-ta.*
      Taro TOP [$_{teP}$ bag   ACC shoulder-GER] parking.lot DAT come-PST

   c. *Taroo wa tyuusyazyoo ni* [$_{teP}$ *kaban o   katui-de*]   *ki-ta.*
      Taro TOP parking.lot DAT [$_{teP}$ bag   ACC shoulder-GER] come-PST

The last variant, (35c), however, also has an interpretation as a V-*te* V complex predicate, because *kuru* 'come' can function as an aux V_2 in the V-*te* V frame. Thus, the surface string in this example is completely ambiguous unless it is pronounced with a differentiating pitch pattern (see Section 3.3).

3.3 Morphophonological unity of V-*te* V

The above considerations lead to the following question: on what grounds should a V-*te* V combination be analyzed as a complex predicate rather than as a mere adjunction structure where a *te*P is adjoined to a main verb. In the cases where the semantics of V_2 is obviously lightened compared with those of their main-verb counterparts, we may be able to safely conclude that they do not contain two full predicates, and thus they do not involve a mere adjunct structure. In other cases such as the V-*te* *kuru/iku* 'come/go' predicates with their physical interpretation retained, however, judgments are more difficult to make on semantic grounds. In regard to this issue, several diagnostics have been proposed in the literature. One such diagnostic is NPI-licensing (McCawley and Momoi 1986; Miyagawa 1987; Matsumoto 1996). By definition, a negative polarity item (NPI) such as *any X* in English and *X-sika* 'anything but X' in Japanese must be locally licensed by a negative head (NEG). In many languages including Japanese, NPI-licensing crossing an adjunct-clause boundary is impossible. This constraint also applies to a regular adjunct *te*P:

(36) a. *Taroo wa* [$_{teP}$ **nanimo** *katuga-***nai**-*de*]   *tyuusyazyoo ni   ki-ta.*
    Taro   TOP [$_{teP}$ anything shoulder-NEG GER] parking.lot   DAT come-PST
    'Taro came to the parking lot without shouldering anything.'

   b. **Taroo wa* [$_{teP}$ **nanimo** *katui-de*]   *tyuusyazyoo ni   ko-***nakat**-*ta.*
    Taro   TOP [$_{teP}$ anything shoulder-GER] parking.lot   DAT come-NEG PST

However, when V_1-*te* in (36) is concatenated to a proper V_2, the NPI is successfully licensed crossing the *te*P boundary:

(36) c. *Taroo wa tyuusyazyoo ni   **nanimo** [katui-de   ko]-***nakat**-*ta.*
    Taro   TOP parking.lot   DAT anything [shoulder-GER come]-NEG PST
    'Taro came to the parking lot without shouldering anything.'

The acceptability of (36c) can be easily accounted for if we assume that V_1-*te* and V_2 jointly form a single predicate in syntax, nullifying the adjunct-clause boundary.

Another convenient test is the "crossed scrambling" test (Miyagawa 1987; Matsumoto 1996). When VP$_1$-*te* is merely adjoined to V_2 without forming a complex

predicate, the internal arguments of these two verbs cannot be crossed by scrambling (i.e. they must be at least nested):

(37) a. *Taroo wa* **tyuusyazyoo ni** [**kaban o** *katui-de*] *mukat-ta.*
　　　 Taro　TOP　parking.lot　　DAT　[bag　　ACC shoulder-GER] head-PST
　　　 'Taro headed for the parking lot shouldering a bag.'

　　 b. **Taroo wa* [**kaban o**　**tyuusyazyoo ni** *katui-de*]　　*mukat-ta.*
　　　 Taro　　TOP　[bag　　　ACC parking.lot　　DAT shoulder-GER] head-PST

By contrast, if the second verb is a qualified aux V, such crossed scrambling is acceptable:

(38) a. *Taroo wa* **tyuusyazyoo ni**　**kaban o**　[*katui-de*　　*it*]-*ta.*
　　　 Taro　TOP　parking.lot　　DAT bag　　　ACC [shoulder-GER go]-PST
　　　 'Taro went to the parking lot shouldering a bag.'

　　 b. *Taroo wa* **kaban o**　**tyuusyazyoo ni**　[*katui-de*　　*it*]-*ta.*
　　　 Taro　　TOP　bag　　ACC parking.lot　　DAT [shoulder-GER go]-PST
　　　 'Taro went to the parking lot shouldering a bag.'

This indicates that the argument structures of the two verbs are "unified" at some level in the V-*te* V constructions. These two tests are basically applicable to most instances of V-*te* V predicates. Other tests to illustrate the morphological unity of V-*te* V complexes are suggested in the literature, such as the long passive test (Matsumoto 1991, 1996), the nominative/accusative alternation tests with the potential morpheme *-re* (Miyagawa 1987), and with the desiderative morpheme *-tai* (Miyagawa 1987; Matsumoto 1996), and a reading time measure (Nakatani 2006).

　　It is worth noting that concatenated V-*te* V predicates may show a distinct pitch pattern from their non-concatenated counterparts. For example, the sequence of *tabe-te* 'eat-GER' and *it-ta* 'go-PST' can be pronounced with the following pitch patterns in accord with the two possible interpretations (Shibatani 2007: 109):[6]

(39) a. Regular adjunct: [*tabe-te*] [*it*]-*ta*
　　　　　　　　　　　 H L L　　L H　 (L=Low pitch, H=High pitch)

　　 b. Concatenated:　 [*tabe-te it*]-*ta*
　　　　　　　　　　　 H L L　　L L

---

6 Shibatani (2007: 109) puts LHHHH as the pitch pattern for (39b), which we assume to be a typo for HLLLL.

The pitch pattern shown in (39a), with two high moras, is a mere sequence of the pitch patterns of *tabe-te* and of *it-ta*. By contrast, the pitch pattern in (39b) "manifests the pitch pattern of a single word" (Shibatani 2007: 109), because (at least in the Tokyo dialect) only one stretch of high-pitched moras is allowed in one word. Shibatani states that the absence of the final pitch shift in (39b) indicates that the sequence forms a "single phonological word."

It is also known that *-te* and the following aux V_2 may optionally be contracted when the initial segment of V_2 is a vowel (Martin 1975: Chapter 9 [Yumoto, this volume] and Chapter 10 [Sugioka and Ito, this volume]). The general rules are as follows. Of the two adjacent vowels, namely the final vowel of *-te* and the initial vowel of V_2, the higher vowel can be omitted. When the two adjacent vowels are of the same height, the [+front] vowel may be omitted:

(40) a. *-te ageru* 'give' → *-t'ageru*

b. *-te iku* 'go' → *-te'ku*

c. *-te iru* 'exist' → *-te'ru*

d. *-te oku* 'put' → *-t'oku*

An exception for this rule is *-te aru* 'exist (inanimate)', which does not contract (**-t'aru*, but *-taaru* is possible in the Osaka dialect). In addition, the sequence of *-te simau* 'put away' can be contracted into *-tyau* or *-timau*, even though the onset of V_2 is not a vowel.[7] These contraction facts further support the assumption that the V-*te* V predicate forms a phonologically tight unit.

3.4 Capturing the duality of V-*te* V

The discussions so far revealed the syntactically ambivalent status of V-*te* V complexes: They behave like a single predicate with regard to NPI licensing, cross-scrambling, pitch pattern, phonological contraction, and other phenomena (Section 3.3), but at the same time they exhibit properties associated with a larger phrasal structure with respect to intervention of focus markers and passivization/causativization of the verbs in V_1 (Section 3.1). How can this ambivalence be accounted for? In the literature, a number of attempts have been made to capture this dual character. Perhaps the most popular analyses are those utilizing "predicate-raising" coupled with tree-pruning (Harada 1977; Kuroda 1978; McCawley and Momoi 1986; Shibatani

---

7 One of the contracted forms, *tyau*, has wider usage (especially among younger speakers) than non contracted *te simau*. For example, an example such as *Kisu s are tyat ta* 'I was kissed' can be uttered either regretfully or delightfully, but if *te simat ta* or the other contracted form *timat ta* were used in place of *tyat ta*, only the regretful connotation would emerge (Yanai 2009: 26). This may indicate that *tyau* is departing from *te simau*, transforming into an independent gramma ticalized form.

1978; Miyagawa 1987), in which the duality is captured in terms of the complex deep structure and its simplified surface structure. An unorthodox analysis is McCawley and Momoi's (1986) "double-mother" analysis, in which the duality is captured by postulating two nodes that dominates a single V_1 head (which is the daughter of the projection of V_1 and of the projection of V_2 at the same time). Still another account is based on LFG (Matsumoto 1996), in which the duality is captured in terms of the discrepancy between the constituent structure (two morphological words) and the functional/argument structures (one word).

None of these analyses, however, address the nature of *-te* itself: they all treat it as a morphosyntactically and semantically vacuous element. Such treatment is problematic because *-te* does seem to contribute to the interpretation of V-*te* V complexes. Consider, for example, the contrast between the V-*te* V predicate and the V-*ni* V predicate, another type of complex predicate that is often called the "purposive motion predicate" (Miyagawa 1987; Matsumoto 1991, 1996):

(41) a. *Sono gakusei wa manga sika yon-de ko-nakat-ta.*
   that student TOP comic only.NPI read-GER come-NEG PST
   'That student read nothing but comic books before coming here.'

  b. *Sono gakusei wa manga sika yomi ni ko-nakat-ta.*
   that student TOP comic only.NPI read to come-NEG PST
   'That student came to read nothing but comic books.'

Superficially, (41a) and (41b) are minimally different: the only difference is found in the choice of the embedding marker: *-te* (realized as *-de* here because of the preceding nasal; see footnote 5) in the former and *ni* in the latter. Yet the temporal interpretations of the V_1 and V_2 are sharply contrastive: in the former, the V_1 event is completive relative to the V_2 event, while it follows the V_2 event in the latter. This contrast suggests that these two small elements, *-te* and *ni*, are not vacuous. They bear some aspectual/temporal functions.

It should also be noted that, as mentioned in (34) in Section 3.2, *-te* in general can host causative and/or passive structures, and this capability is retained in the V-*te* V predicate. It is a standard assumption in generative grammar that causative and passive morphemes select verbal projections (which include internal arguments if any) rather than a verb-stem itself. Because *-te* is placed even outside such a projection, it is natural to assume that *-te* selects a phrasal category larger than VP, *v*P, AspP, and even NegP.

These and other considerations motivate a head movement analysis, which is a refined and updated version of the predicate-raising analyses in the early generative framework. Under this analysis (Nakatani 2013), *-te* is regarded as a tense/aspectual marker (let us assume it is T) that denotes a relative past, and a V-*te* V complex is formed through several stages of successive head movements, as illustrated in (42).

(42) a. $[_{vP/VP} \ldots V_1]$ (Embedded VP)

    b. $[_{TP} [_{vP/VP} \ldots V_1]\text{-}te]$ (*-te* Merges)

    c. $[_{TP} [_{vP/VP} \ldots t] V_1\text{-}te]$ (V$_1$ internally Merges with *-te*)

    d $[_{VP} [_{TP} [_{vP/VP} \ldots t] V_1\text{-}te] V_2]$ (*te*P Merges with V$_2$)

    e. $[_{VP} [_{TP} [_{vP/VP} \ldots t] t] V_1\text{-}te\text{-}V_2]$ (V$_1$-*te* internally Merges with V$_2$)

    f. $[_{VP} (\text{NP-DAT}) [_{TP} [_{vP/VP} \ldots t] t] V_1\text{-}te\text{-}V_2]$ (Dative argument, if any, Merges)

The constituency of VP$_1$-*te* is present at earlier stages in the course of derivation, namely (42b–d). A focus marker may be attached to it at stage (42c). The movement (namely, internal Merge) of V$_1$ to *-te* is driven by a morphological factor (*-te* is a bound morpheme). The movement of V$_1$-*te* to V$_2$, on the other hand, is not morphologically motivated, because V$_2$ is a free morpheme. Rather, the movement must be semantically motivated, enabled by a special predicate-concatenation mechanism, which only a handful of verbs can trigger. The output, i.e. the V$_1$-*te*-V$_2$ complex, is thus interpreted in a special way; as a result, the complex may or may not license a dative argument. This means that this dative argument, if any, is an argument of the V$_1$-*te*-V$_2$ complex, not just an argument of V$_2$. This assumption accounts for the various valency patterns found in the V-*te* V constructions (Section 4.2). This head movement approach also explains the NPI-licensing fact and the possibility of cross-scrambling, as well as phonological properties of the predicate (pitch patterns, contraction).

3.5 Problems of the embedded subject

Finally, let us touch on one of the most perplexed issues regarding the syntax of V-*te* V complex predicates: the status of the embedded subject (for V$_1$) distinct from the matrix subject (for V$_2$). There are three constructions, namely V-*te morau*, *-te hosii*, and *-te aru*, that involve a "distinct" embedded subject, as illustrated below.

(43) a. V-*te morau* ('be given')

      Watasi$_i$ wa kare ni   e$_i$ home-te   morat-ta.

      I$_i$ TOP he   DAT *e$_i$* praise-GER be.given PST

      'I was praised by him (I benefited from him praising me).'

    b. *Watasi wa kare ni ie   ni ki-te   morat-ta.*

      I   TOP he   DAT house DAT come-GER be.given PST

      'I had him come to my house (I benefited from him coming to my house).'

(44) a. V-*te hosi-i* ('want')

   Watasi$_i$ wa kare ni e$_i$ home-te hosi-i.
   I$_i$ TOP he   DAT e$_i$ praise-GER want PRS
   'I want him to praise me.'

   b. *Watasi wa kare ni ie   ni ki-te   hosi-i.*
   I   TOP he   DAT house DAT come-GER want PRS
   'I want him to come to my house.'

   c. *Watasi wa ame ga   hut-te   hosi-i.*
   I   TOP rain NOM fall-GER want PRS
   'I want it to rain.' (lit. 'I want rain to fall.')

(45) a. V-*te ar-u* ('exist')

   *Kyookasyo$_i$ ga   { e$_j$ / *watasi ni   / *watasi ga }   e$_i$ kat-te   ar-u.*
   textbook$_i$ NOM { e$_j$ / *I   DAT / *I   NOM } e$_i$ buy-GER exist PRS
   'The textbook has been bought (I have bought the textbook).'

   b. *Watasi$_j$ wa kyookasyo$_i$ o   kat-te   ar-u.*
   I$_j$ TOP textbook$_i$ ACC buy-GER exist PRS
   'I have bought the textbook.'

The (a) sentences in the above examples all show passive-like structures, where the matrix subject is identified with the object of the embedded verb V$_1$. In (43b) and (44b, c), on the other hand, there is no direct thematic relationship between the matrix subject and the embedded predicate V$_1$. The "embedded subject" is dative-marked in V-*te morau* (43) and V-*te hosii* (44), which is reminiscent of the dative subjects attested in passives and causatives. As for V-*te hosii*, an option of marking the embedded subject with nominative case is also available, as shown in (44c) (see also Section 2.3). This is a peculiar property not found in the other V-*te* V constructions.

The syntactic properties of V-*te aru* complexes are even more peculiar, in that the construction does not allow the embedded subject to be overt in any form as shown in (45a), and that it also allows the *in situ* structure as in (45b). The former is sometimes called in the literature the "intransitivizing" (Martin 1975: 524; Matsumoto 1990) or "passive-type" -*te aru* constructions (Masuoka 1987: 220). As mentioned earlier, what is peculiar here is that the embedded subject (e$_j$) is obligatorily null in (45a), meaning that having a dative subject, which is widely attested in various constructions in Japanese, is not an available option here. For this reason, Matsumoto (1990) argues that the intransitivizing V-*te aru* construction involves no embedded subject position in syntax. This issue is also discussed by Sugioka (1986: 185–189) and Nakatani (2013: 159–163).

4 The semantics of V-*te* V predicates

We now move on to consider the semantics of V-*te* V predicates. It is very well known that the meanings of aux V_2s are often lightened (bleached), which is one of the motivations for calling them auxiliary verbs. A deeper question to ask is whether there are general rules operative across the semantic lightening effects found in various instances of the V-*te* V predicates. Yet another valid question is how they are captured in the linguistic knowledge in a native speaker's mind. Although we do not reject application of the label "auxiliary verb" to the V_2s in the V-*te* V constructions, we argue that they are not completely fixed as such in the native speaker's knowledge. Rather, it can be assumed that the native speakers have in their mind some "theories" regarding the auxiliary status of V_2s in the V-*te* V constructions relative to their main-verb counterparts. The evidence for this hypothesis is as follows. First, as pointed out in Section 1, the aux V_2s and their main-verb counterparts coexist in modern Japanese, and neither is anywhere near being obsolete. Second, the morphosyntactic properties of the main-verb V_2s are completely retained in the aux V_2s, although there are some differences observed in their phonological properties, as discussed in Section 3.3. Third, the suppletive honorification forms for the main-verb V_2s (46a) may be used in the V-*te* V predicate (46b) (if not always).

(46) a. *Sensei   ga   { i   / irassyat   / orare* }-*ta*.
teacher NOM { exist / exist.HON / exist.HON }-PST
'The teacher was present.'

   b. *Sensei   ga   ne-te   { i   / irassyat   / orare* }-*ta*.
teacher NOM sleep-GER { exist / exist.HON / exist.HON }-PST
'The teacher was sleeping.'

Fourth, the semantic interpretations of aux V_2 are often determined interactively with the semantics of VP_1 and *-te*. This last fact is particularly prominent in the case of V-*te kuru/iku* ('come/go'). It is known that *kuru* and *iku* in the V-*te* V frame show a wide variety of interpretations, but the point is that they are mostly determined in relation to the semantics of VP_1. In other words, the properties of *kuru* and *iku* are dependent on the syntactic context in which they are used. This is what Pustejovsky (1995) calls "complementary polysemy". Pustejovsky discusses in detail the general problems of dealing with this type of polysemy by enumerating different meanings; he argues that such cases should be treated as involving context-dependent sense extension. We discuss this issue in Section 4.1. Finally, an association between the aux V_2s and their main-verb counterparts would enable us to explore general rules applicable across various V-*te* V constructions, as suggested later in Section 4.2.

4.1 The polysemy of *kuru* 'come' and *iku* 'go'

As mentioned above, V-*te kuru/iku* 'come/go' complexes provide an interesting set of data ranging from usage very close to its main verb semantics (physical movement) to use as a more abstract, aspectual marker. Previous studies, mostly done in the fields of traditional Japanese linguistics, grammaticalization, and cognitive linguistics, focus on identifying distinct semantics of the construction and setting the borderline between the "auxiliary" and "full verb" usage of *-te kuru/iku* (Masuoka 1992; Morita 1968, 1994; Yoshikawa 1973; Moriyama 1988; Teramura 1984; Yamamoto 2000). For example, Teramura (1984: 156–163) summarizes the semantic relations between V_1-*te* and V_2 as follows:

(47) a. Type V-V (Semantically coordinative)
   Tukare-ta   kara,   tyotto   koohii o   non-de   ku-ru.
   get.tired-PST because, one.moment coffee ACC drink-GER come-PRS
   'I got tired, so I'll go drink a cup of coffee (drink coffee and come back).'

   b. Type v-V (V_1 is semantically subordinate)
   Gakkoo e   wa zitensya ni   not-te   ki-mas-u.
   school toward TOP bicycle DAT ride-GER come-POL PRS
   'I go to school by bicycle.'

   c. Type V-v (V_2 is semantically light)
   Sora ga   akaruku nat-te   ki-ta.
   sky NOM bright   become-GER come-PST
   'The sky started to grow light.'

Example (47a) contains two sequential events, and thus the V-*te* V predicate represents coordination-like semantics. In (47b), on the other hand, V_1 denotes the manner of the V_2 event, and in this sense, V_1 is semantically subordinate (hence v-V). In (47c), V_2 has lost the sense of physical movement, and thus, it is auxiliary-like (hence V-v). Teramura (1984) identifies V_2 in Type V-v as an aspectual marker, and distinguishes it from the others. Teramura (1984: 158) suggests two criteria for identifying Type V-v (although the second criterion seems to be merely an elaboration of the first):

(48) a. *X ga … V-te ku-ru* [X NOM … V-GER come-PRS] does not entail *X ga ku-ru* [X NOM come] 'X comes';

   b. *X ga … V-te ku-ru* [X NOM … V-GER come-PRS] does not denote X's physical movement; rather, it denotes that the situation represented by *X ga … V* approaches the speaker.

For example, *kuru* 'to come' by itself can be predicated of its subject in (47a, b) (i.e. the speaker actually "comes"), whereas it cannot in (47c) (i.e. the sky does not come

to us). Teramura (1984) argues that *kuru* in the latter is a "secondary" constituent that adds the sense of the event denoted by V_1 coming toward the speaker (more on this in Section 4.3). See also Shibatani (2007) for detailed discussion from the perspective of grammaticalization, and Nakatani (2015) for a criticism of Shibatani's approach.

It should be emphasized that Teramura's classification of aux *kuru/iku* (as well as Shibatani's "cline of grammaticalization") is highly context-dependent. That is to say, one particular type found in a specific context cannot be replaced by the other types. Thus, *kuru* in (47a) cannot be interpreted as an aspectual marker, and *kuru* in (47c) cannot be interpreted as denoting a physical motion. This is a case of "complementary polysemy" (Pustejovsky 1995), for which a sense enumeration approach cannot provide an adequate explanation. Rather, a derivational mechanism should be postulated to account for the distribution of various types of *kuru/iku* depending on the context (i.e. the VP$_1$-*te* that *kuru/iku* is attached to).

4.2 Valency and the principle of causation flow

As obvious from the above discussion, and as explicitly pointed out in the literature (Hamada 1989; Alam 1999; Yamamoto 2000; Shibatani 2007; Nakatani 2003b, 2013), the varied meanings of aux *kuru* and *iku* are often predictable from their relationship with the semantics of VP$_1$. Empirical evidence for this view comes from the availability of a goal argument in V-*te kuru/iku* predicates (Nakatani 2003b, 2013; Shibatani 2007). Although *kuru* 'come' and *iku* 'go' themselves are capable of licensing a goal argument marked with the locative particle *ni*, the concatenated V-*te* V frame may or may not license such arguments. Compare the four sentences in (49) with special attention to the goal argument *kaisya ni* 'to his office'.

(49) a. *Kare wa kaban o* ***kaisya ni*** *kakae-te* *it-ta.*
      he   TOP bag   ACC office   LOC hold.under.arm-GER go PST
      'He went to his office, holding his bag under his arm.'

   b. *Kare wa kaban o* (****kaisya ni***) *kakae-ta.*
      he   TOP bag   ACC (*office   LOC) hold.under.arm-PST
      'He held his bag under his arm (*to his office).'

   c. *Kare wa sirase o* (****kaisya ni***) *kii-te* *it-ta.*
      he   TOP news ACC (office   LOC) hear-GER go PST
      'He heard the news and (then) went (*to his office).'

   d. *Kare wa sirase o* *kii-te* ***kaisya ni*** *it-ta.*
      he   TOP news ACC hear-GER office   LOC go PST
      'He heard the news and (then) went to his office.'

The V-*te iku* complex in (49a) allows the goal argument 'to his office' to appear on the immediate left of the V1 *kakae-te* 'holding', which by itself is unable to license such a goal argument, as shown by the ungrammaticality of (49b). On the other hand, the V-*te* V sequence in (49c), *kii-te itta* [hear-GER went], cannot take the goal phrase 'to his office', even though the motion verb *itta* 'went' alone can, as shown in (49d). How can the discrepancy between *kakae-te it-ta* 'went holding' (49a) and *kii-te it-ta* 'went hearing' (49d) be accounted for?

The problem of the goal argument *kaisya ni* 'to his office' in (49c) is that although it is certainly an argument of *iku* 'go', it does not fit well as an argument of a larger phrase [*sirase o kii-te iku*] 'hear the news and go' as a whole. On the other hand, the same phrase functions perfectly as a goal argument of the [*kaban o kakae-te iku*] 'hold the bag and go' complex in (49a), which invokes a "carrying/bringing" interpretation as a whole. This intuition – *kaisya ni* 'to his office' is a legitimate argument only if it fits well as an argument of the complex predicate as a whole – is confirmed by the fact that such a goal argument is perfectly acceptable if *iku* is not concatenated to V$_1$ *kii-te* as in (49d). Put differently, the problem about (49c) emerges because the relationship between the two internal arguments of the complex, *sirase o* 'the news' and *kaisya ni* 'to the office', are expected to constitute a "coherent" argument structure but actually turns out to be somewhat disjoint.

As discussed earlier in Section 3.3, V-*te* V complexes allow NPI-licensing across the *-te* boundary, as well as crossed scrambling of the internal arguments of V$_1$ and V$_2$. This fact suggests that the V-*te* V complexes form a single predicate at some level, which in turn indicates that they may have the two argument structures of V1 and V2 unified into one.[8] Adopting the idea of a prototypical causal chain proposed in

---

8 One may consider that if this "unified argument structure" hypothesis is true, long passive (Matsumoto, 1991, 1996), i.e. passivization targeting the whole V *te* V complex, should always be allowed, which is actually not the case: some allow long passive, others do not:

(i) *Kare   wa   okusan   o   ture te   ki ta.*
     He   TOP   wife   ACC   took GER   come PST
     'He brought his wife with him.'

(i') *Okusan   ga   ture te   ko rare ta.*
     wife   NOM   took GER   come PASS-PST
     'His wife was brought here.'

(ii) *Kare   wa   okusan   o   home te   oi ta.*
     He   TOP   wife   ACC   praise GER   put PST
     'He praised his wife (in advance).'

(ii') **Okusan   ga   home te   ok are ta.*
     wife   NOM   praise GER   put PASS-PST
     (Intended: 'His wife was praised (in advance).')

The reason why (ii') is unacceptable is that V$_2$ *ok u* represents a certain attitude or intention of the logical subject *kare* 'he' ("he did so in advance so that it would yield a desirable outcome later"), which is irrelevant to the state of the object of V$_1$, *okusan* 'wife'. When the sentence is passivized as in (ii'), the object *okusan* is promoted as the grammatical subject, which clashes with the logical subject oriented function of *ok u*. This is just one instance, and other cases call for further examination.

the literature (such as "Argument Coherence" and "Default Causative Paradigm" in Pustejovsky and Busa 1995 and Pustejovsky 1995, "Direct Object Restriction" in Levin and Rappaport Hovav 1995, and "Billiard-ball Model" in Langacker 1987), Nakatani (2003b, 2013) attempts to capture the observed contrast in terms of the interpretive well-formedness condition imposed on the syntactically overt arguments that is grounded on the "principle of causation flow". Simply put, the principle of default, canonical causation flow as stated in (50) below is enforced if the result event is made "syntactically explicit" by the presence of a syntactically overt argument that is exclusively associated with the result event (Nakatani 2013: 188–191, 261).

(50) If the causing event involves a theme, the result event should depict what happens to the theme (rather than to the agent).

Kaisya ni 'to the office' cannot be overt in (49c) because it would make the result event syntactically explicit, enforcing the "causation flow" in (50) strictly. Such a canonical causal interpretation is unavailable: rather, it means 'he' heard the news (causing event), and then 'he' came (result event). The sentence would be acceptable if the goal argument were covert, which would shadow (i.e. background) the result event; in such a case, the canonical causal interpretation in (50) is not enforced. On the other hand, the dative argument in (49a) is acceptable because the theme *kaban* 'bag' can be construed as reaching the goal place *kaisya* 'office' by virtue of the implication from VP$_1$ 'holding the bag'. This interpretation fits the canonical causation flow: 'he' held the 'bag' in his arms (causing event), and then the 'bag' reached the site (result event). Nakatani (2014) further argues that this principle is responsible for the availability of the goal arguments in the V-*te kureru/ageru* 'give (to the speaker/a non-speaker)' constructions as well.

4.3 Theme and events

It is frequently observed that the semantics of the aux V$_2$ is lightened to the extent that an abstract event rather than a physical object is taken as the theme argument, a process that is widely attested in grammaticalization phenomena across many languages (e.g. Heine 1993; Hopper and Traugott 1993). For example, the aspectual sentence in (47c), repeated below, cannot be interpreted in such a way that the sky itself came to the speaker, as depicted in (51).

(47) c. Type V-v (V$_2$ is semantically light)
   Sora   ga   akaruku   nat-te   ki-ta.
   sky   NOM   bright   become-GER   come-PST
   'The sky became increasingly bright.'

(51)
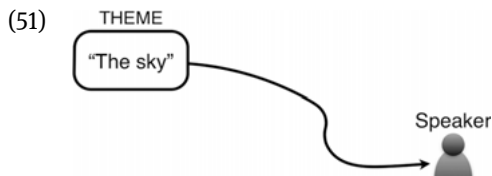
Instead, many researchers (Ikegami 1981; Teramura 1984; Masuoka 1992; Yui 1997; among others), assume that in this kind of example, the whole event represented by *X ga … V₁* comes into the perspective of the speaker (see (48b)), as schematically represented in (52).

(52)
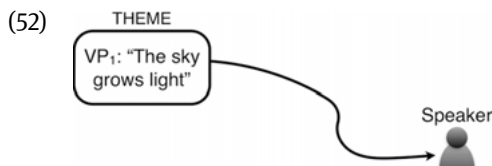
In this interpretation, the denotation of the theme argument is shifted from the domain of physical objects to that of events. It should be emphasized here that the goal concept involved in the original V₂ semantics is *not* lost in the V₁-*te* V₂ predicate, as evident in the contrast shown in (3), repeated below:

(3) a. *Zutuu   ga   si-te   { ki-ta   / *it-ta }.*
    headache NOM do-GER {come-PST / *go-PST}
    'I am getting a headache.'

   b. *Syoonen wa dandan   tuyoku nat-te   { ki-ta   / it-ta }.*
    boy   TOP gradually strong become-GER {come-PST / go-PST }
    'The boy was getting stronger and stronger.'

As discussed in Section 2.1, the change in the speaker's condition (the headache) depicted in (3a) is most naturally described as "coming" toward the speaker, and hence only *kita* 'came' is acceptable but not *itta* 'went'. In (3b), on the other hand, the change in the third party *syoonen* 'boy', on the other hand, can be perceived as either "coming" to the speaker's recognition or "going" up the stages, so both versions are acceptable. These considerations lead us to assume that even those aux V₂s retain a goal component in their semantic structure, which in turn necessitates a theme component because whenever there is a goal, there should be something moving toward the goal. For this reason, we would continue to use the term "theme" even when no physical object undergoes a movement.

   This analysis can be extended to the other instances of V-*te* V predicates discussed in Section 2. In fact, such a "domain shift" of the theme is generally obligatory in the V-*te* V predicates except for the V-*te kuru/iku* constructions (see (47a, b) and (49a, b)). Other (seemingly) exceptional cases are illustrated in (53)–(55) below.

(53) *Kanozyo ga obentoo o tukut-te* { *kure-ta* / *age-ta* }.
she   NOM lunch.box ACC make-GER { give-PST / give-PST }
'She fixed a lunch box (for the speaker/for someone).'

(54) *Kanozyo ga obentoo o teeburu ni tukut-te oi-ta*.
she   NOM lunch.box ACC table   LOC make-GER put-PST
'She fixed a lunch box on the table.'

cf. **Kanozyo ga obentoo o teeburu ni tukut-ta*.
she   NOM lunch.box ACC table   DAT make-PST

(55) *Tonari no heya ni kodomo tati ga ne-te i-ta*.
next   GEN room LOC child   PL   NOM sleep-GER exist-PST
'The children were sleeping in the room next door.'

cf. **Tonari no heya ni kodomo tati ga ne-ta*.
next   GEN room LOC child   PL   NOM sleep-PST

In (53), the "object sharing" interpretation is available, where the object *obentoo* 'lunch box' can be interpreted as the arguments of both V$_1$ and V$_2$ (i.e. 'she fixed a lunch box and gave it to the speaker or to someone else'). However, it may be the case that this object sharing interpretation is made available by pragmatic inference, because the activity of fixing a lunch box, which is supposed to be beneficial by virtue of the semantics of *kureta/ageta* ('give'), implies that the lunch box is given to the recipient. In (54) and (55), on the other hand, the "physical theme" interpretation is certainly necessary because of the presence of the locative arguments, *teeburu ni* 'to the table' and *tonari no heya ni* 'in the next room'. Such examples, however, are not widely attested, and the majority of the V-*te oku* and V-*te iru* examples are not compatible with locative arguments that are not licensed by V$_1$. Thus, in the V-*te* V constructions except for a subset of the V-*te kuru/iku* predicates, the "domain-shift" of V$_2$'s theme from physical objects to events is the norm, not the exception.

Under the "domain-shift" approach discussed so far, the event denoted by VP$_1$ is identified as the theme, as illustrated in (52). This would give us an impression that aux V$_2$ directly takes VP$_1$ as its internal argument. However, a closer examination would reveal that such an analysis is not sufficiently accurate. This point has already been discussed in Section 2.3 in relation to the semantics of V-*te miru*: it does *not* mean that the subject "sees" the VP$_1$ event itself; rather it means that the subject "sees" what has come out of the VP$_1$ event (see (18)–(20)). The same point can be observed in other V-*te* V complexes. For example:

(56) *Taroo ga denwa o kake-te ki-ta*.
Taro NOM telephone ACC place-GER come-PST
'Taro called (me).'

(57) *Hanako ga   tomato o   tabe-te kure-ta.*
   Hanako NOM tomato ACC eat-GER give-PST
   'Hanako ate the tomatoes for me.'

(58) *Hanako wa syukudai   o   yat-te { oi-ta   / simat-ta }.*
   Hanako TOP homework ACC do-GER { put-PST / put.away-PST }
   'Hanako has finished the homework.'

Under the prominent interpretation of (56), Taro does not make a physical move-ment: he places a call or sends a postcard, staying where he is. Thus, the theme for V$_2$ *kita* cannot be *denwa* 'telephone'. Then, is it the VP$_1$ event? If so, the denotation of, say, (56) would be informally illustrated as follows:

(59)
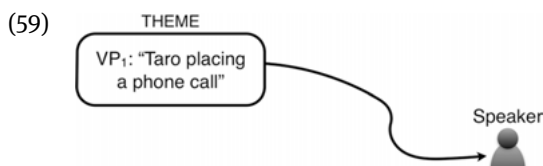
This, however, does not really capture what the sentence really means. Rather, it is the result of his placing a call (such as phone-ringing or telephone message) that is metaphorically directed toward the speaker.

In order to resolve this issue, we need several assumptions concerning the domain-shifted theme (Nakatani 2008, 2013: 194–197). First, let us assume that *-te* functions as a relative past tense marker (event sequencer), and what is directed toward the speaker is not the VP$_1$ event itself, but an event logically or pragmatically interpreted as some sort of continuation or consequence of the VP$_1$ event, an event made available by "stretching" the VP$_1$ event, borrowing from Landman's (1992) notion of "stage and continuation stretch":

(60)   An event is a stage of another event if the second can be regarded as a more developed version of the first. (Landman 1992: 23)

(61)   The continuation stretch of an event *e* is the smallest set of events *f* such that *e* is a stage of *f*. (Adopted from Landman 1992: 26)

In this analysis, the semantics of (56) can be computed as in (62), where *-te* is a relative past tense marker, and a logically or pragmatically inferred "stretch" of the VP$_1$ event is taken to be the theme of V$_2$.

(62)
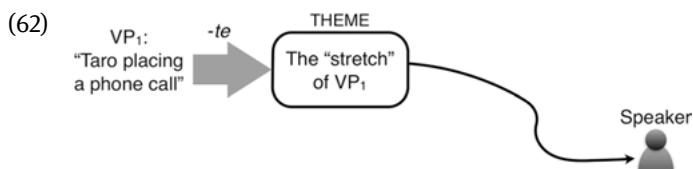
In this case, an inferred continuation/consequence of Taro's activity of placing a phone call is the phone call itself (or the event of the phone ringing), and it is this phone call (or the phone-ringing event) that is reaching the speaker. Nakatani (2013) further argues that the "stretched" event may result either in a resultative reading or in a continuative reading, depending on the aspectual properties of VP$_1$. This is particularly notable in the well-known aspectual contrast found in the interpretation of V-*te iru*.

(63) a. *Tori ga   sin-de   i-ru.*
  bird NOM die-GER exist-PRS
  'The bird is dead.' (perfective)

 b. *Tori ga   ton-de i-ru.*
  bird NOM fly-GER exist-PRS
  'The bird is flying.' (progressive)

Here the aspectual property of VP$_1$ influences the aspectual interpretation of the whole complex predicate. *Sinu* 'die' is telic, and hence a resultative reading is strongly preferred in computing a stretch of it, resulting in the perfective interpretation of *sin-de iru* in (63a). *Tobu* 'fly', on the other hand, is an atelic activity predicate, and thus, a continuation of it is preferably (if not decisively; see the discussion on (7) in Section 2.1) taken to be the stretch of it.

 Let us consider the semantics of the V-*te oku/simau* constructions given in (58). The *oku* and *simau* versions denote the same events, and are indistinguishable truth-conditionally, but the former implies that the subject, *Hanako*, is ready for the next move, while the latter emphasizes the completion of the homework, indicating that she does not have to worry about it and she is now free. Under the present framework (see also Ono (1992), who suggests essentially the same analysis), this contrast can be accounted for in the following way. First, we can assume that the full verbs *oku* and *simau* both denote an event of the agent putting the theme onto some location. In the case of *oku*, the location is a two-dimensional surface, and in the case of *simau*, it is a three-dimensional closed space. In the V-*te oku/simau* constructions, on the other hand, the theme undergoes a domain-shift, and the implied consequence of the temporally preceding VP$_1$ event, in this case, the completion of the homework, is taken to be the theme that is placed onto a surface or into a closed space. Note that an object placed onto a surface is generally easy to access, whereas it is difficult to access when placed into a closed container. Thus, depending on the choice between *oku* and *simau*, different connotations emerge: the state resulting from the completion of the homework is either ready to access (*oku*) or placed beyond the reach (*simau*). This analysis nicely explains the emergence of regretful connotations in some V-*te simau* constructions, as illustrated in (64):

(64) *Kabin o   wat-te   simat-ta!*
     vase ACC break-GER put.away-PST
     'I broke the vase!'

A regretful connotation emerges in the above example because the consequence of the VP$_1$ event, which is undesirable, is placed beyond the reach of the speaker, implying that it cannot be undone (Ono 1992). The use of *oita* in place of *simatta* in this example would be very awkward. By contrast, the VP$_1$ event in (58), i.e. doing the homework, is something that one would hate to deal with, and thus, no regretful nuance emerges.

In closing this section, special attention should be paid to the semantics of the two aux verbs of existence *iru* and *aru*.

(65) *Taroo wa syukudai o   yat-te* { *i-ru   / ar-u* }.
     Taro TOP homework ACC do-GER { exist PRS / exist-PRS }
     'Taro { is doing the homework / has done the homework }.'

In (65), the *iru* version denotes a progressive event and the *aru* version a perfective result. The latter can be straightforwardly accounted for in terms of the domain-shift of the theme: Taro did the homework, and the consequence inferred (or stretched) from this event is "there", at an unspecified discourse space (Masuoka 1987; Matsumoto 1990). It is known that V-*te aru* complexes with volitional subjects are usually accompanied with a sense of "preparation for a future task" like the one found with V-*te oku* predicates (Yoshikawa 1973; Takahashi 1976; Masuoka 1987; Matsumoto 1990). This is not surprising because, if the consequence is "there", it should be easy to access, just as something put on a surface is accessible. The volitionality requirement for VP$_1$ in the V-*te ar-u* constructions (Takahashi 1976; Masuoka 1987; Tsujimura 1991), as seen from the ungrammaticality of (66) below, most likely stems from this connotation: the "future task" reading forced by the semantics of *aru* requires the VP$_1$ event to involve a volitional agent.

(66) **Yuki ga   Okinawa ni   hut-te   ar-u.*
      snow NOM Okinawa LOC fall-GER exist-PRS
      'There was a snowfall in Okinawa.'

The V-*te iru* construction is more problematic to our analysis. We cannot apply the same domain-shift account to it, because the semantics of the *iru* construction would otherwise be identical to that of the *aru* construction. One possibility is to hypothesize that while V-*te aru* involves domain-shift of the theme, V-*te iru* does not (Nakatani 2013: 227–230). This hypothesis is partly motivated by the fact that the V-*te aru* predicate is generally more lenient in allowing locative arguments, as illustrated by the sharp contrast between (67a) and (67b) (Hasegawa 1992).

(67) a. *Taroo wa* (?*soko ni*) *syukudai o   yat-te ar-u.*
      Taro   TOP there LOC homework ACC do-GER exist PRS
      'Taro has done the homework (which can be found over there).'

    b. *Taroo wa* (*soko ni) *syukudai o   yat-te i-ru.*
      Taro   TOP there LOC homework ACC do-GER exist PRS
      'Taro is doing the homework (*which can be found over there).'

While V-*te aru* in (67a) is marginally acceptable on the reading where the "consequence" of finishing the homework is visible in the place designated by *soko* 'there', V-*te iru* in (67b) totally lacks such an interpretation. This suggests that the theme of *iru* has not really undergone domain-shift. Finally, the requirement for the "preparation" connotation found in V-*te aru* constructions is completely absent in V-*te iru* constructions. For example, the predicates that are inherently incompatible with the "preparation for the future" sense, such as *sinu* 'die', can enter the V-*te iru* frame but not the V-*te aru* frame, indicating that the sense of the eventive theme being "there" is absent in the V-*te iru* predicate.

(68) a. *Tori ga   sin-de i-ru.*
      bird NOM die-GER exist-PRS
      'The bird is dead.'

    b. **Tori ga   sin-de ar-u.*
      bird NOM die-GER exist-PRS

5 Conclusion and future research perspectives

This chapter has examined twelve representative instances of V-*te* V complex predicates, focusing on the associative rules and principles that enable us to make predictions regarding the syntactic and semantic properties of these varied constructions. Although the V_2s in the complexes look like auxiliaries from the viewpoint of their lightened meanings, it has been shown that many of the properties are predictable from the semantics of their main-verb counterparts and their relation to the VP_1 semantics. Based on these considerations as well as the fact that the aux V_2s and their main-verb counterparts synchronically coexist, we have argued that many of the seemingly grammaticalized properties are mentally computed synchronically in accordance with the inputs (either in the course of acquisition or in the adult grammar, or both). The combined structure of V_1, *-te*, and V_2 has been analyzed as a syntactically concatenated predicate. Issues of embedded subjects have been discussed in this light. Issues of valency and semantic bleaching effects have been analyzed in terms of canonical causation flow and the theme's domain shift, respectively. Under the "domain-shift" analysis, aux V_2 may take an event as its theme, but

it has been argued that it is *not* the VP$_1$ event itself that becomes the theme of V$_2$ (hence VP$_1$ is *not* the complement of V$_2$); rather, VP$_1$'s pragmatically inferred "stretch" (either a consequence or continuation of the VP$_1$ event) is interpreted as V$_2$'s theme.

What we can see from all these considerations is that the V-*te* V constructions show a considerable number of predictable regularities, although some unpredictable idiosyncrasies remain. Open questions include the syntax of -*te aru* (especially the questions of why the dative subject is never allowed and why the object of V$_1$ has options of staying *in situ* or of being promoted to the subject), the syntax and semantics of -*te hosii* (especially the status of *te*P in this construction), the differences between -*te kuru* vs. -*te iku* that cannot be simply reduced to the directionality contrast (see, for example, footnotes 1 and 3), the problem of "hostile" -*te yaru* (footnote 4) and "joyful" -*tyau* (footnote 7), and the problem of the morphosyntactic complexity, or the "size", of the V-*te* V constructions in general (they are definitely larger than lexical V-V compounds, and smaller than adjunct-gerund constructions, but are they larger than syntactic V-V compounds? See footnote 8 for a brief discussion on the applicability of long passive). Some of these problems may turn out to be idiosyncratic to specific instances of the V-*te* V constructions (thus calling for a grammaticalization perspective), while others may be resolved under a more general, derivational theory. Future research is expected to carefully tease apart the predictable and unpredictable aspects of the syntax and semantics of the V-*te* V constructions so that we are not left with a mere collection of random observations. Any theory of the lexicon should be a plausible model of the mental knowledge of native speakers, which is likely to comprise regularities and idiosyncrasies.

Acknowledgements

References

Alam, Sachiko Sasaki. 1999. Idō hōkō no hatsuwa jōkyō izonsei [Directionality of motion and its dependency on discourse]. In Sachiko Sasaki Alam (ed.), *Gengogaku to nihongo kyōiku* [Linguistics and Japanese pedagogy], 177–196. Tokyo: Kurosio Publishers.

Hamada, Mariko. 1989. "Iku/kuru" to "teiku/tekuru" no imi no tsunagari [*Iku / kuru* and *te iku / te kuru* in Japanese]. *Sophia Linguistica* 27. 47–56. Sophia University.

Harada, S. I. 1977. The derivation of unlike-subject desideratives in Japanese. *Descriptive and Applied Linguistics* 10. 131–146.

Hasegawa, Yoko. 1992. On the ambiguity between the perfect and the resultative: V-*te ar* constructions in Japanese. In Laura A. Buszard-Welcher, Lionel Wee and William Weigel (eds.), *Proceedings of the eighteenth annual meeting of the Berkeley Linguistics Society*, 88–99. Berkeley: Berkeley Linguistics Society.

Hasegawa, Yoko. 1996. *A study of Japanese clause linkage: The connective* te *in Japanese*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. Oxford: Oxford University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.

Ikegami, Yoshihiko. 1981. *"Suru" to "naru" no gengogaku* [The linguistics of "make" and "become."]. Tokyo: Taishukan Shoten.

Imani, Ikumi. 1990. V-*tekuru* to V-*teiku* ni tsuite [On V-*tekuru* and V-*teiku*]. *Nihongogaku* 5. 54–66.

Jacobsen, Wesley. this volume. Chapter 15: Lexical meaning and temporal aspect. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. this volume. Chapter 8: Verb-compounding and verb-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Komatsu, Toshio. 1964. *Shite yaru, shite morau, shite kureru* [-Shite yaru, -shite morau, -shite kureru]. In Kenji Morioka, Ken Nagano, Yutaka Miyaji and Takashi Ichikawa (eds.), *Yurete iru bunpō* [Swaying grammar], 193–202. Tokyo: Meiji Shoin.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuroda, S.-Y. 1978. Case marking, canonical sentence patterns, and counter equi in Japanese. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics*, 30–51. Tokyo: Kaitakusha.

Landman, Fred. 1992. The progressive. *Natural Language Semantics* 1. 1–32.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar, volume I: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax lexical semantics interface*. Cambridge, MA: MIT Press.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Masuoka, Takashi. 1987. *Meidai no bunpō* [Grammar of proposition]. Tokyo: Kurosio Publishers.

Masuoka, Takashi. 1992. Nihongo no hojo dōshi kōbun [Auxiliary verb constructions in Japanese]. In *Bunka gengogaku: Sono teigen to kensetsu* [Cultural linguistics: Its proposal and setup], 546–532. Tokyo: Sanseido.

Matsumoto, Yo. 1990. On the syntax of Japanese "intransitivizing" -*te aru* construction. In Michael Ziolkowski, Manuela Noske and Karen Deaton (eds.), *Papers from the 26th Regional Meeting of the Chicago Linguistic Society, Volume 1: The Main Session*, 277–292. Chicago, IL: Chicago Linguistic Society.

Matsumoto, Yo. 1991. On the lexical nature of purposive and participial complex motion predicates in Japanese. In Laurel A. Sutton, Christopher Johnson and Ruth Shields (eds.), *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, 180–191. Berkeley: Berkeley Linguistics Society.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Matsuo, Sutejirō. 1936. *Kokugohō ronkō* [Considerations on Japanese grammar]. Tokyo: Hakuteisha.

Matsushita, Daizaburō. 1930. *Kaisen hyōjun nihon bunpō* [Revised standard Japanese grammar]. Tokyo: Chūbunkan.

McCawley, James D. and Katsuhiko Momoi. 1986. The constituent structure of *-te* complements. *Papers in Japanese Linguistics* 11. 1–60.

Miyagawa, Shigeru. 1987. Restructuring in Japanese. In Takashi Imai and Mamoru Saito (eds.), *Issues in Japanese linguistics*, 273–300. Dordrecht: Foris.

Miyake, Tomohiro. 2005. Gendai nihongo ni okeru bunpōka [Grammaticalization in modern Japanese]. *Nihongo no Kenkyū* 1. 61–75.

Morita, Yoshiyuki. 1968. *Iku / Kuru* no yōhō [The usage of *iku / kuru*]. *Kokugogaku* 75. 75–87.

Morita, Yoshiyuki. 1977. *Kiso Nihongo* [Basic Japanese]. Tokyo: Kadokawa Shoten.

Morita, Yoshiyuki. 1994. *Dōshi no imironteki bunpō kenkyū* [A semantic and grammatical study of verbs]. Tokyo: Meiji Shoin.

Moriyama, Takuro. 1988. *Nihongo dōshi jutsugobun no kenkyū* [A study of verbal predicates in Japanese]. Tokyo: Meiji Shoin.

Muraki, Masatake. 1978. The *sika nai* construction and predicate restructuring. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics*, 155–177. Tokyo: Kaitakusya.

Nakatani, Kentaro. 2003a. Analyzing *-te*. In William McClure (ed.), *Japanese/Korean Linguistics 12*, 377–387. Stanford: CSLI Publications.

Nakatani, Kentaro. 2003b. V-concatenation in Japanese: A GL approach. In Pierrette Bouillon and Kyoko Kanzaki (eds.), *Proceedings of the second international workshop on generative approaches to the lexicon*, 107–114.

Nakatani, Kentaro. 2006. Processing complexity of complex predicates: A case study in Japanese. *Linguistic Inquiry* 37. 625–647.

Nakatani, Kentaro. 2008. *Tekuru*, *teiku* no dōshi kyōki seigen no hasei [Deriving the verb collocation constraints on *tekuru* and *teiku*]. In Taro Kageyama (ed.), *Lexicon forum* 4. 63–89. Tokyo: Hituzi Syobo.

Nakatani, Kentaro. 2013. *Predicate concatenation: A study of the V te V predicate in Japanese*. Tokyo: Kurosio Publishers.

Nakatani, Kentaro. 2014. Shieki-rensa no gensoku to *te*-kei fukuzatsu-jutsugo ni okeru *ni*-kaku no yōninsei [Principle of causation flow and acceptability of *ni*-dative in the V-*te* V complex predicate]. In Hideki Kishimoto and Yoko Yumoto (eds.), *Fukuzatsu jutsugo kenkyū no genzai* [Current trends in the studies of complex predicates], 99–124. Tokyo: Hituzi Syobo.

Nakatani, Kentaro. 2015. *Te*-kei fukuzatsu-jutsugo no tagisei o dō toraeru beki ka: Bunpōka apurōchi to kakudai gōsei apurōchi [On how the polysemy in the V-*te* V predicate should be dealt with: Grammaticalization approach and enriched composition approach]. *The Journal of Konan University: Faculty of Letters* 165. 99–112. Konan University.

Nishida, Naotoshi. 1977. Joshi (1) [Particles (1)]. In Susumu Ōno and Takeshi Sibata (eds.), *Iwanami kōza nihongo* 7: Bunpō II [Iwanami Japanese linguitstics series 7: Grammar II]. Tokyo: Iwanami Shoten.

Ono, Tsuyoshi. 1992. The grammaticization of the Japanese verbs *oku* and *shimau*. *Cognitive Linguistics* 3(4). 367–390.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pustejovsky, James and Federica Busa. 1995. Unaccusativity and event composition. In Pier M. Bertinetto, Valentina Bianchi and James Higginbotham (eds.), *Temporal reference: Aspect and actionality, volume 1: Semantic and syntactic perspectives*, 159–177. Torino: Rosenberg and Sellier.

Shibatani, Masayoshi. 1978. *Nihongo no bunseki: Seisei bunpō no hōhō* [An analysis of Japanese: The method of generative grammar]. Tokyo: Taishukan Shoten.

Shibatani, Masayoshi. 2007. Grammaticalization of motion verbs. In Bjarke Frellesvig, Masayoshi Shibatani and John Charles Smith (eds.), *Current issues in the history and structure of Japanese*, 107–133. Tokyo: Kurosio Publishers.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Takahashi, Taro. 1976. Sugata to mokuromi [Aspect and intention]. In Haruhiko Kindaichi (ed.), *Nihongo dōshi no asupekuto* [Aspect of Japanese verbs], 117–153. Tokyo: Mugi Shobō.

Teramura, Hideo. 1984. *Nihongo no shintakusu to imi II* [The syntax and semantics of Japanese II]. Tokyo: Kurosio Publishers.

Tsujimura, Natsuko. 1991. On the semantic properties of unaccusativity. *Journal of Japanese Linguistics* 13. 91–116.

Yamada, Toshihiro. 2004. *Nihongo no benefakutibu* [Benefactives in Japanese]. Tokyo: Meiji Shoin.

Yamamoto, Hiroko. 2000. "Kuru" no tagi kōzō [Polysemic structure of *kuru*]. *Nihongo Kyōiku* 105. 11–20.

Yanai, Hisae. 2009. *Teshimau* sōtō keishiki no imi kinō kakuchō [Semantic-functional expansion of *-te shimau* and its variants from the Edo period to present]. *Nihongo no Kenkyū* 5. 15–30.

Yoshikawa, Taketoki. 1973. Gendai nihongo no asupekuto no kenkyū [A study of aspect in Modern Japanese]. *Linguistic Communications* 9. Monash University. Reprinted in Haruhiko Kindaichi (ed.), 1976, *Nihongo dōshi no asupekuto* [Aspect of Japanese verbs], 155–327. Tokyo: Mugi Shobō.

Yui, Kikuko. 1997. Nihongo dōshi ni okeru imi no chūshō-ka katei no kenkyū: Hojo dōshi yōhō o motsu dōshi no imi bunseki [A study of the process of abstraction of Japanese verbs: Analyses of meanings of verbs possessing auxiliary usages]. *Osaka Daigaku Bungakubu Kiyō* 37. 1–152. Osaka University.

Tadao Miyamoto and Hideki Kishimoto
12 Light verb constructions with verbal nouns

1 Introduction

Broadly speaking, verbs function as the predicate of a sentence while nouns function as arguments. This basic division of labor, however, is sometimes blurred. One such case is a "light verb construction" (LVC), in which the verb basically carries only tense and the argument structure information necessary to interpret the whole sentence is furnished by the nominal that seemingly functions as the verb's object, as illustrated by the English examples in (1).

(1) a. *Tom took a look at the lady.*
     b. *Tom gave the corridor a sweep.*
     c. *Tom had a drink of sake.*

In (1), *take*, *give*, and *have*, termed "light verbs" by Jespersen (1966), do not express their literal meanings but instead serve as predicates in combination with the nouns *look*, *sweep*, and *drink*. Light verbs can thus be thought of as verbs that have little semantic content of their own and express fully predicative meanings in alliance with certain additional nouns denoting activities or events. The term "light verb construction" refers to such combinations of a light verb and a certain kind of nominal that as a whole function as the predicate of a sentence. Since the substantial predicate meanings come from the object noun phrases (*a look*, *a sweep*, *a drink*), light verb constructions can in principle be paraphrased by replacing the light verbs with inflected verbs related to the nominals, as in (2).

(2) a. *Tom looked at the lady.*
     b. *Tom swept the corridor.*
     c. *Tom drank sake.*

Jespersen's original idea has been extended and applied to similar constructions in many languages other than English. Grimshaw and Mester (1988) identify constructions like those in (3) as the Japanese counterparts of light verb constructions.

(3) a. *Taroo   ga   Hanako   to   AISEKI   o   si-ta.*
         Taro   NOM   Hanako   with   table.sharing   ACC   do-PST
         'Taro shared a table with Hanako.'

     b. *Taroo   ga   rai-getu   Kyooto   ni   RYOKOO   o   su-ru.*
         Taro   NOM   next-month   Kyoto   to   travel   ACC   do-PRS
         'Taro takes a trip to Kyoto next month.'

In (3a) and (3b), the native verb *suru* 'do', which appears to have little semantic content, is identified as a Japanese light verb. The predicative contents of (3) are attributed to the nouns AISEKI 'table-sharing' and RYOKOO 'travel', which are marked in the accusative (*o*). These nouns are identified as verbal nouns (VNs) rather than simple nouns because they have a predicative function with argument structure (here and elsewhere in this chapter, VNs will be represented in small capitals). Specifically, the comitative phrase (*to* 'with') in (3a) and the goal phrase (*ni* 'to') in (3b) are selected by AISEKI 'share a table (with)' and RYOKOO 'travel (to)', respectively, in the same way that the preposition *at* in (1a) is selected by the noun *look* rather than by the light verb *take*. The function of *suru* is to make it possible for VNs, which do not inflect by themselves, to exhibit not only tense (VN *o su-ru* (non-past), VN *o si-ta* (past)) but also aspect and modality (e.g. VN *o si-sooda* 'be about to do VN') as well as passive (VN *o s-are-ru*) and causative (VN *o s-ase-ru*).

A special morphological property of the Japanese constructions in (3) is that the accusative case marker on VNs is optionally omitted, and if it is omitted, a VN and *suru* take on the appearance of a compound word (VN-*suru*) (see Chapter 14 [Kageyama, this volume] for the morphological status of VN-*suru*). As a result, the compound-verb constructions in (4a) and (4b) below express exactly the same meanings as the periphrastic constructions in (3a) and (3b), respectively.

(4) a. *Taroo ga   Hanako to   AISEKI-si-ta.*
      Taro   NOM Hanako with table.sharing-do-PST
      (the same meaning as 3a)

   b. *Taroo ga   rai-getu   Kyooto ni RYOKOO-su-ru.*
      Taro   NOM next-month Kyoto   to trip-do-PRS
      (the same meaning as 3b)

In this chapter, we will restrict the term "light verb constructions" to the periphrastic constructions of type (3) and discuss their semantic and syntactic properties. Section 2 will survey a variety of uses of *suru* to identify the nature of the light verb *suru*. Section 3 will enumerate the syntactic characteristics of the light verb constructions. Section 4 will discuss some special properties of the VN *o suru* construction, and review Grimshaw and Mester's (1988) seminal work and subsequent studies with particular reference to the so-called light *suru* hypothesis and heavy *suru* hypothesis. Section 5 is a sample corpus study to fathom the nature of the VN *o suru* construction from the viewpoint of Miyamoto (1999). Section 6 provides perspectives for future research.

2 Various uses of the verb *suru*

The verb *suru* 'do' is used in a huge variety of constructions. This section will outline the various uses of *suru* with a view to clarifying the nature of *suru* as a light verb.

2.1 Transitive and intransitive uses of *suru*

Like the English verb *do*, the verb *suru* functions prototypically as a two-place predicate designating a volitional action or activity of the agent (subject), e.g. *warui koto o suru* [bad thing ACC do] 'do bad things', *arubaito o suru* [part.time.job ACC do] 'do a part-time job', *okurimono o suru* [gift ACC do] 'give a gift', but it also has a number of distinct uses that are not found with the English *do*. Table 1 illustrates a range of the transitive uses of *suru* (cf. Kageyama 1993).

Table 1: Uses of transitive *suru*

	examples	meaning
A. Game playing	*{tenisu/sakkaa/toranpu/booringu} o suru* {tennis/soccer/cards/bowling} ACC do	'play {tennis/soccer/cards/bowls}'
B. Professions	*{isya/kyoosi/bengosi} o si te iru* {doctor/teacher/lawyer} ACC do GER be	'be a {doctor/teacher/lawyer}'
C. Resultative	*Taroo wa musuko o {isya/bengoshi} ni si ta.* Taro TOP son ACC {doctor/lawyer} DAT do PST	'Taro made his son a {doctor/lawyer}.'
D. Innate attributes	*Zyon wa aoi me o si te i ru.* John TOP blue eye ACC do GER be PRS	'John has blue eyes.'
E. Wearing accessories	*Taroo wa akai nekutai o si te i ru.* Taro TOP red necktie ACC do GER be PRS	'Taro is wearing a red tie.'

The *suru* in A takes the names of games as its objects, whereas that in B selects object nouns denoting professions in the auxiliary construction with *-te iru*. The construction in C, which is regarded as the causative or resultative counterpart of B, is licensed only in the construction with an accusative object and a dative complement (result phrase) meaning 'to make somebody something' or 'to designate something to something' as in *Taroo wa senkoo o koogaku ni si-ta.* [Taro TOP major ACC engineering DAT do-PST] 'Taro decided on engineering as his major'. Whereas the verb *suru* in A, B, and C designates volitionally controllable actions, the construction in D expresses non-volitional inalienable possession.

Two prominent restrictions are imposed on the D construction (Kageyama 2004). Observe the examples in (5).

(5) a. **Tomu   wa   aoi   me   o   {su-ru/si-ta}.*
         Tom   TOP   blue   eye   ACC   {do-PRS/do-PST}
         'Tom {has/had} blue eyes.'

     b. *[aoi   me   o   {si-ta/si-te   iru}]   ningyoo*
         blue   eye   ACC   {do-PST/do-GER   be-PRS}   doll
         'a doll with blue eyes'

c. *Merii   wa   kawaii   hokuro   o   si-te   i-ru.
   Mary   TOP   cute   mole   ACC   do-GER   be-PRS
   'Mary has a cute mole.'

In the first place, the sentence form of the D construction is syntactically restricted to the one where the body-part expression is followed by the complex form *si-te iru* expressing the state of the subject, so that the plain past or present form, as in (5a), is prohibited. But the simple past form *si-ta*, as well as *si-te iru*, is allowed when the D construction is used as an attributive modifier, as in (5b). In the second place, the range of the accusative-marked nouns is semantically limited to body parts that a human is endowed with innately, such as eyes, hands, legs, etc.; other bodily features like *hokuro* 'mole' are excluded, as shown by the unacceptability of (5c).

As opposed to the "innate attribute" construction in D which represents the subject's permanent or individual-level property, the construction in E expresses a temporary or stage-level state of 'wearing an alienable item on one's body'. Object nouns in the latter construction are limited to those nouns denoting accessories like a necktie, a scarf, and a ring that one wears for ornament, and *boosi* 'hat', *kutu* 'shoes', *uwagi* 'jacket', and the like are not qualified, as the unacceptability of (6b) shows.

(6) a. Hanako   wa   sutekina   {nekutai/sukaahu/yubiwa}   o   si-te   i-ru.
       Hanako   TOP   nice   {necktie/scarf/ring}   ACC   do-GER   be-PRS
       'Hanako wears a nice {necktie/scarf/ring}.'

    b. *Taroo   wa   sutekina   {boosi/kutu/uwagi}   o   si-te   i-ru.
       Taro   TOP   nice   {hat/shoes/jacket}   ACC   do-GER   be-PRS
       'Taro wears a nice {necktie/scarf/ring}.'

According to Kageyama (1980), the impossibility of *suru* 'do' in (6b) is due to lexical blocking (or pre-emption) by three special verbs that are dedicated to clothing: *kaburu* 'wear on one's head' must be used for a hat, a cap, etc.; *haku* 'wear on one's legs/feet' for shoes, pants, socks, etc.; and *kiru* 'wear on one's body or trunk' for a jacket, a shirt, etc. As a consequence, the verb *suru* can take care of only those ornamental items (a necktie, a ring, etc.) to which the three dedicated verbs do not apply. (For accessories, it is possible to use other verbs, e.g. *nekutai o simeru* [necktie ACC tie] 'tie a necktie', *yubiwa o hameru* [ring ACC put.on] 'wear a ring', *iyaringu o tukeru* [earring ACC attach] 'wear earrings'.)

The verb *suru* also has intransitive uses as summarized in Table 2.

Table 2: Intransitive uses of *suru*

	examples	meaning
F. Price	*Kono tokei wa zyuuman en si ta.* this watch TOP 100,000-yen do-PST	'This watch cost 100,000 yen.'
G. Perception	*{Oisisoona nioi/ookina oto} ga si ta.* {delicious smell/big sound} NOM do-PST	'There was {a delicious smell/a big sound}.'

The construction in F designates the price or cost of goods. In the construction of G, a perceived bodily sensation is expressed with the nominative subject representing a stimulus of perception such as smell, sound, taste, and internal feeling. For some reason or other, visual perception like color is excluded, as shown by the ungrammaticality of (7).

(7) **Azayakana iro ga su-ru.*
      bright   color   NOM   do-PRS
      'There is a bright color.'

The transitive-intransitive distinction of *suru* may involve lexical idiosyncrasies that are at best attributed to the semantics of the subject or object nouns in individual examples. While the two sentences in (8) both involve physiological phenomena, (8a) with the noun *geri* 'diarrhea' is expressed transitively, and (8b) with the nouns *memai* 'dizziness' and *ikigire* 'the state of being out of breath' is represented intransitively.

(8) a. *Taroo wa geri   o   si-ta.*
          Taro   TOP diarrhea   ACC do-PST
          'Taro suffered from diarrhea.'

     b. *Taroo wa {memai/ikigire}   ga   si-ta.*
          Taro   TOP {dizziness/out.of.breath} NOM   do-PST
          'Taro {felt giddy/got out of breath}.'

There are also lexical restrictions on argument nouns. For example, the transitive construction in (8a) is at odds with such diseases as *haien* 'pneumonia' and *hara-ita* 'stomach ache'. On the other hand, the intransitive *suru* can also take nouns of natural happenings such as *amamori* 'rain leakage' and *inabikari* 'lightning', as in (9).

(9) *{Amamori/Inabikari}   ga   si-ta.*
      {rain.leakage/lightning} NOM   do-PST
      'There was {a leak of rain/lightning}.'

It is plausible to conclude that the intransitive *suru* takes as its subject those nouns that represent physical sensations, physiological states, and natural phenomena – in other words, nouns that denote events which take place spontaneously without external causers.

2.2 The light verb *suru*

So far, we have reviewed the variety of uses of *suru* outside the domain of light verb constructions. This discussion is intended to highlight the special nature of the light verb *suru* in the "VN *o suru*" construction. In all the non-light verb constructions, the range of co-occurring nouns, which are regarded as plain nouns that lack argument structures, is heavily constrained to individual lexical items. On the other hand, the light verb *suru* does not impose any idiosyncratic lexical restriction on its co-occurring nouns but only specifies the lexical category of VN.

The VNs used in the light verb constructions are mostly Sino-Japanese (e.g. *KOOGI* 'lecture', *RYOKOO* 'travel') as in (10a), but may also be native Japanese (e.g. *AZIMI* 'taste', *KASI KARI* 'lending and borrowing', *TATI YOMI* 'browsing') as in (10b) and Western (e.g. *DORAIBU* 'driving', *COPII* 'copy') as in (10c) (see Chapter 2 [Kishimoto and Uehara, this volume]).

(10) a. *Kyoozyu   wa   daigaku   de   gengogaku no   KOOGI   o   si-ta.*
professor TOP university LOC linguistics GEN lecture ACC do-PST
'The professor gave a lecture on linguistics at the university.'

b. *Taroo   wa   hon-ya   de   zassi   no   TATI YOMI   o   si-ta.*
Taro   TOP book-store LOC magazines GEN stand-reading ACC do-PST
'Taro browsed magazines in a bookstore.'

c. *Taroo   ga   mai-syuu   Hakone made   DORAIBU o   su-ru.*
Taro   NOM every-week Hakone as.far.as driving   ACC do-PRS
'Taro drives as far as Hakone every week.'

Syntactically, VNs may be marked in the accusative case as superficial objects of the verb *suru*, as in (10), or may be combined with *suru* by leaving out the accusative marker as in (11).

(11) a. *Kyoozyu   wa   daigaku   de   gengogaku o   KOOGI-si-ta.*
professor TOP university LOC linguistics ACC lecture-do-PST
(the same meaning as 10a)

b. *Taroo   wa   hon-ya   de   zassi   o   TATI YOMI-si-ta.*
Taro   TOP book-store LOC magazines ACC stand-reading-do-PST
(the same meaning as 10b)

c. *Taroo ga mai-syuu Hakone made DORAIBU-su-ru.*
Taro NOM every-week Hakone as.far.as driving-do-PRS
(the same meaning as 10c)

Compounding with *suru* as in (11) is a privilege allowed only to VNs. None of the plain nouns in the analogous constructions in Tables 1 and 2 are allowed to be compounded with *suru*. Apparent counter-examples are native nouns like *ikigire* 'the state of being out of breath' and *kiokure* 'feeling of being daunted' (Hirao 1995), which may or may not take the nominative case in the intransitive use of *suru*, as shown in (12a) and (12b).

(12) a. *ikigire ga su-ru*   b. *ikigire-su-ru*
     out.of.breath NOM do-PRS   out.of.breath-do-PRS
     'be out of breath'   'be out of breath'

However, *ikigire* and *kiokure* do not seem to have argument structure and they in fact do not behave like VNs in other respects. Specifically, full-fledged VNs have a special ability to function as predicates without the support of *suru* when they are put in certain types of tenseless subordinate clauses discussed by Shibatani and Kageyama (1988) and elaborated on by Kageyama (1993). The sentences in (13) are thus fully grammatical and convey the same contents as (10).

(13) a. *Kyoozyu ga daigaku de gengogaku o KOOGI no sai …*
     professor NOM university LOC linguistics ACC lecture GEN time
     'When the professor gave a lecture on linguistics at the university, …'

   b. *Taroo ga hon-ya de zassi o TATI YOMI no sai …*
     Taro NOM book-store LOC magazines ACC stand-reading GEN time
     'When Taro browsed magazines in a bookstore, …'

   c. *Taroo ga Hakone made DORAIBU no sai …*
     Taro NOM Hakone as.far.as driving GEN time
     'When Taro drove to Hakone, …'

The nouns *ikigire* and *kiokure*, however, do not fit in with this construction, as in (14).

(14) **Taroo ga {ikigire/kiokure} no sai …*
     Taro NOM {be.out.of.breath/feel.daunted} GEN time
     'when Taro {got out of breath/felt daunted}, …'

The drop of the case particle on these nouns might be attributed to a separate rule of N-V compounding (cf. Chapter 7 [Kageyama, this volume]). The same remark holds of

cases like *otya-suru* [tea-do] 'go out for tea' and *seisyun-suru* [youth-do] 'enjoy young days', where non-VNs (which do not have their own argument structure) are combined with *suru*.

We will conclude this section by calling attention to the use of *suru* as a "dummy" verb (Kageyama 1992), which should be strictly distinguished from the light verb *suru*. Observe the examples in (15).

(15) a. *Taroo wa Hakone made* DORAIBU *(o) si sae si-ta.*
    Taro   TOP Hakone up.to drive   (ACC) do even do-PST
    'Taro even drove his car to Hakone.'

   b. *Taroo wa sono hon o kai wa si-ta ga, yom-anakat-ta.*
    Taro   TOP the book ACC buy TOP do-PST but read-NEG PST
    'Taro only bought the book, and didn't read it.'

(15a) displays two occurrences of *si-*, the *ren'yō* form of *suru*. Since the first *si-* is the light verb associated with the VN DORAIBU 'drive', the second *si-* cannot be a light verb. Instead, the latter *si-* serves as a dummy verb supporting a stranded (tense) morpheme, which is separated from the main verb by a particle like *sae* 'even' and *wa* [TOP]. In fact, this dummy verb can also occur after a plain native verb like *kai* 'buy' in (15b), and it is in complementary distribution with another dummy verb *aru* 'be' which occurs in the same environment where tense is separated from adjectives and ANs in the presence of intervening particles.

3 Syntactic realizations of light verb constructions

Let us now turn to discussion of the syntactic characteristics of verbal nouns (VNs) and the light verb *suru*. In this section, we will first consider some notable properties of verbal nouns, and then discuss how VN-*suru* and VN *o suru* should be derived.

3.1 Duality of VN forms

VNs are nouns possessing certain verbal behaviors: they serve as predicates when combined directly with *suru*, and act as nominal arguments when they bear case marking (Martin 1975). In the Japanese lexicon, there are a large number of VNs, which are either native Japanese words (e.g. YAMA NOBORI 'mountain-climbing'), Sino-Japanese words (e.g. SOODAN 'consulting') or foreign words borrowed from such languages as English (e.g. KISU 'kissing'). Incidentally, mimetic words can occur in front of *suru*, as in *Atama ga hurahura suru* [head NOM (MIM) do] 'I feel giddy', but this construction generally does not count as a light verb construction,

because mimetics do not function as predicates without the support of *suru* (**Atama ga hurahura no sai* [head NOM (MIM) GEN time], as is the case with *ikigire-suru* 'out of breath' discussed above (cf. (14)); see also Chapter 4 [Akita and Tsujimura, this volume]). Among the native, Sino-Japanese, and foreign strata, the Sino-Japanese stratum is by far the richest source of VNs, so in what follows we will focus on Sino-Japanese VNs and discuss their properties.

Observe at this point that most Sino-Japanese VNs are composed of two bound morphemes. While the two morphemes are arranged in diverse relations, as discussed in Chapter 3 (Kobayashi, Yamashita and Kageyama, this volume), three representative types of internal composition may be recognized, as seen in (16).

(16) a. V-V: *TOO KAI* 倒壊 [fall-destroy] 'collapsing'

b. ADV-V: *SOO TAI* 早退 [early-leave] 'early leaving'

c. V-N: *DOKU SYO* 読書 [read-book] 'reading'

In (16a), one verbal morpheme is compounded with another verbal morpheme (i.e. *too-kai* ≈ *taore-te kowareru* [fall-GER destroy] 'fall and collapse'); in (16b), a verbal morpheme is preceded by an adverbial morpheme (i.e. *soo-tai* ≈ *hayaku sirizoku* [early leave] 'leave early'); and in (16c), a verbal morpheme is followed by a nominal element that corresponds to the internal argument of the verb (i.e. *dokusyo* ≈ *syo o yomu* [book ACC read] 'read books') (see Kageyama 1980; Jacobsen 1991). The morpheme orders in these three classes are obviously due to the word orders of Chinese.

The light verb VN *o suru* form has a VN-*suru* counterpart, as in *KEKKON o suru* [marry ACC do] versus *KEKKON-suru* [marry-do] 'to get married'. VN-*suru* forms are generally regarded as compound words, while VN *o suru* forms are undoubtedly syntactic phrases comprising two separate elements. The two types of light verb forms give rise to a difference in acceptability when a VN expands to a phrase.

(17) a. *[suugaku no BENKYOO] o su-ru*
     [math   GEN study]   ACC do-PRS
     'to study mathematics'

b. **[suugaku no BENKYOO]-su-ru*
     [math   GEN study]-do-PRS
     'to study mathematics'

When a VN is directly combined with the verb *suru*, the entire form looks like a single word. The following gapping example provided by Kageyama (1999) further shows that VN-*suru* behaves like a single word.

(18) a. *_Tuma wa_ DAIETTO-s̶i̶, _otto wa_ KIN'EN-_si-ta._
      wife TOP diet-do   husband TOP quit.smoking-do-PST
      'The wife went on a diet, and the husband quit smoking.'

    b. _Tuma wa_ DAIETTO _o_ -s̶i̶, _otto wa_ KIN'EN _o si-ta._
      wife TOP diet   ACC do husband TOP quit.smoking ACC do-PST
      'The wife went on a diet, and the husband quit smoking.'

As seen in (18a), gapping is prohibited from deleting only the _suru_ of a VN-_suru_ form, which should be possible if the VN and _suru_ were two separate constituents syntactically. In fact, the gapping of _suru_ is fully legitimate if the VN is accompanied by the accusative marker _o_ and makes up a separate constituent from _suru_, as in (18b). This fact suggests that the VN-_suru_ behaves as a complex word.

    Nevertheless, the VNs in the VN-_suru_ form sometimes show behavior as a phrase. There are a number of indications that VN and _suru_ also function as separate elements syntactically. One such indication is provided by the fact that VNs can often be coordinated to the exclusion of the verb _suru_.

(19) _Sensei wa_ SANSEI _mo_ HANTAI _mo si-nakat-ta._
    teacher TOP approve also disapprove also do-NEG PST
    'The teacher neither approved nor disapproved.'

In (19), two VNs SANSEI 'approve' and HANTAI 'disapprove' are coordinated, and the verb _suru_ is detached from it. This example thus suggests that a VN and _suru_ can stand separately in syntactic structure. Notice, however, that the syntactic separation is made possible by the presence of _mo_ 'also' and other focus particles. More generally, a VN can be detached from _suru_ if it is accompanied with a focus particle, as in (20).

(20) _Sensei wa_ SANSEI _wa si-ta keredo, atode...._
    teacher TOP approve TOP do-PST but   later ...
    'The teacher expressed his approval, but later ...'

What is important here is that Japanese has a general rule that a case marker is dropped out in the presence of a focus particle. In both (19) and (20), it is entirely plausible to suppose that the accusative case marker on the VNs, though present syntactically, is phonologically deleted in the presence of the focus particles _mo_ and _wa_. If so, sentences like (19) and (20) are variants of the syntactic VN _o suru_ form, not of the compound VN-_suru_ form.

    The plausibility of our assumption is enhanced by the marked contrast between (21a) and (21b), which involve two VNs conjoined by a genuine coordinator _to_ 'and'.

(21) a. *Sensei   wa   [DAIETTO   to   KIN'EN]   o   (ryoohoo)   si-ta.*
     teacher TOP [diet   and quit.smoking] ACC (both)   do-PST
     'The teacher (both) went on a diet and quit smoking.'

   b. **Sensei   wa   [DAIETTO   to   KIN'EN]-si-ta.*
     teacher TOP [diet   and quit.smoking]-do-PST
     'The teacher went on a diet and quit smoking.'

As shown by (21b), coordinated VNs cannot be compounded with *suru* for a reason to be explained shortly.

While the ungrammaticality of (21b) points to the word status of a VN-*suru* form, -*kata* nominalization (Kageyama 1993; Kishimoto 2006) makes the opposite suggestion. Observe how *BENKYOO (o) suru* [study ACC do] 'study' is nominalized by the suffix -*kata* 'way'.

(22) a. Nominalization of VN *o suru*   b.   Nominalization of VN-*suru*
       BENKYOO   no   si-kata   **BENKYOO-si-kata*
       study   GEN do-way   study-do-way
       'the way of studying'   'the way of studying'

The VN *o suru* form is nominalized in (22a), where the accusative marker is replaced by the genitive marker, in accordance with a general rule of Japanese case marking. Quite unexpectedly, however, if the compound VN-*suru* form is nominalized as such, an ungrammatical form like (22b), **BENKYOO-si-kata* [VN-do-way], results. The unavailability of the form in (22b) suggests that the VN and *suru* should be treated as syntactically independent elements.

Incidentally, verbs such as *nes-suru* 'heat' *ai-suru* 'love' look like N-*suru*. Nevertheless when these verbs are nominalized with the suffixation of −*kata* 'way', they have the forms *nessi-kata* [heat-way] 'the way of heating' and *aisi-kata* [love-way] 'the way of loving', and not **netu no si-kata* [heat GEN do-way] and **ai no si-kata* [love GEN do-way]. This fact suggests that *netu* 'heat' and *ai* 'love' are not VNs, i.e., the verbs *nes-suru* and *ai-suru* with the surface N-*suru* forms are not derived by combining VNs with the light verb *suru* (Kageyama 1993).

3.2 Syntactic issues of light verb constructions

There are a number of theoretical and descriptive issues surrounding VN-*suru* forms. One concerns the question of whether the formation of *suru*-form takes place at the lexical or syntactic level (cf. Sugimura 2012 and the references cited therein). There are mainly two different views on the derivation of VN-*suru* forms, i.e. lexical and syntactic approaches. In the lexical approach (Miyagawa 1987), the verb *suru* is

attached to VNs in the lexicon, but in the syntactic approach (Kageyama 1982, 1993), VN and *suru* appear separately in the syntax and are combined into one word by syntactic incorporation.

The data regarding (17) and (21) suggest that VN-*suru* forms are sometimes not available in certain contexts, signaling that they are generated separately in the syntax and then undergo incorporation into *suru* at the level of syntax.

(23) a. *BENKYOO o suru* → [*BENKYOO*]$_{VN}$-*suru*
        study   ACC do   study-do

     b. [*suugaku no BENKYOO*] *o suru* → *[$_{VNP}$ suugaku no BENKYOO*]-*suru*
        math   GEN study   ACC do   math   GEN study-do

Under Baker's analysis (1988), noun incorporation is head movement in syntax, where a lexical category (X^o) is adjoined to another lexical category (Y^o). Because of this condition, the VN that undergoes incorporation must be a bare VN (lexical category) instead of a phrasal category headed by a VN. Consequently, ungrammatical examples like (17b) = (23b) involving a modified VN phrase *suugaku no BENKYOO* 'study of mathematics' or like (21b) involving a coordinate VN phrase *DAIETTO to KIN'EN* 'go on a diet and quit smoking' are correctly ruled out by the head movement constraint. In the lexical analysis, (17b) and (21b) would be ruled out by the No Phrase Constraint, which states that compound words may not contain a syntactic phrase (cf. Botha 1984).

Let us now turn attention to the realization of the arguments of VNs, in particular how the object argument of VNs is realized. For example, the object of the VN *SOOZI* 'cleaning' is realized as a genitive modifier, as in (24a), but it also has an alternative realization as an accusative-marked object, as in (24b).

(24) a. *Gakusei ga [kyoositu no SOOZI] o si-ta.*
        student NOM classroom GEN cleaning ACC do-PST
        lit. 'The students did the cleaning of the classroom.'

     b. *Gakusei ga kyoositu o SOOZI-si-ta.*
        student NOM classroom ACC cleaning-do-PST
        'The students cleaned the classroom.'

In (24a), the VN *SOOZI* 'cleaning' makes up a phrase together with the modifier *kyoositu no* 'of the classroom' and the VN phrase as a whole is marked with the accusative. In (24b), on the other hand, the VN *SOOZI* 'cleaning' is compounded with *suru* and the NP 'classroom' is realized as the direct object of *SOOZI-suru* 'clean'. Note that the two alternating patterns are not always available. Consider the examples in (25a) and (25b).

(25) a. *Gikai   ga   [hooan no   {SAIKETU/??KAKETU}] o   si-ta.*
      parliament NOM bill   GEN {vote/pass}   ACC do-PST
      'The parliament {voted for/passed} the bill.'

   b. *Gikai   ga   hooan o   {SAIKETU/KAKETU}-si-ta.*
      parliament NOM bill   ACC {vote/pass}-do-PST
      'The parliament {voted for/passed} the bill.'

In (25a) using the NP *no* VN *o suru* construction, SAIKETU 'decide by vote' is impeccable while KAKETU 'pass, approve' is questionable. This disparity disappears in the VN-*suru* construction in (25b), where both VNs are perfectly acceptable. Tanomura (1988) and Kageyama (1993) each make a fairly comprehensive survey on the alternations like (24a, b) or (25a, b).

Focusing on agentivity, Tanomura (1988) observed that the NP *no* VN *o suru* construction is generally possible when the VN is agentive in nature. However, both SAIKETU 'decide by vote' and KAKETU 'pass, approve' appear to be agentive. On the other hand, Kageyama (1991, 1993) points out that the grammaticality of the genitive modifier construction hinges on finer-grained parameters controlling the transitivity of a sentence, such as agentivity, punctuality, affectedness, and others. In (25a), the act designated by SAIKETU 'decide by vote' is an accomplishment eventuality involving a complex sequence of voting and deciding, whereas the act of KAKETU 'pass, approve' is an achievement that takes place instantaneously – presumably because of this semantic property, KAKETU, unlike SAIKETU, has an intransitive use, e.g. *Hooan ga {KAKATU/*SAIKETU}-si-ta*. [bill NOM {pass/*vote}-do-PST] 'The bill {passed/*voted}'. The former has a higher degree of transitivity than the latter, and such semantic properties of VNs contribute to the accusative marking in the NP *no* VN *o suru* construction.

Furthermore, the examples in (26) show that the alternations of the genitive modifier (NP *no*) in the (25a)-type construction and the accusative object (NP *o*) in the (25b)-type construction are not automatic, either.

(26) a. *Hanako ga   [{piano/ensookai} no   RENSYUU] o   si-ta.*
      Hanako NOM {piano/concert}   GEN practice   ACC do-PST
      'Hanako did practice {on the piano/for the concert}.'

   b. *Hanako ga   {piano/*ensookai} o   RENSYUU-si-ta.*
      Hanako NOM {piano/concert}   ACC practice-do-PST
      'Hanako practiced {the piano/*the concert}.'

Since *piano* 'piano' is the argument (object) of the VN RENSYUU 'practice', both constructions in (26a) and (26b) are permitted. However, *ensookai* 'concert' is a mere modifier of the VN in (26a), so it cannot appear as the object of the VN-*suru* in (26b).

The assignment of accusative marking to VNs is constrained by the lexical properties of VNs themselves. Right after Grimshaw and Mester (1988) expressed their view that the accusative marker is assigned indiscriminately to VNs by the light verb *suru*, Dubinsky (1989), Miyagawa (1989), Tsujimura (1990), and Kageyama (1991) independently discovered that transitive and unergative VNs allow the accusative marking whereas unaccusative VNs do not readily do so. (Generally speaking, unergatives represent a volitional or controllable action, and unaccusatives a non-volitional, non-controllable event). Omitting the examples of transitive VNs as they have been amply exemplified in the preceding discussion, we present examples of only unergative and unaccusative VNs. Observe the contrast between (27a) and (27b).

(27) a. *Sensei ga {UNDOO-si-ta/UNDOO o si-ta}.*
teacher NOM {exercise-do-PST/exercise ACC do-PST}
'The teacher did some exercises.'

b. *Mizu ga {ZYOOHATU-si-ta/?*ZYOOHATU o si-ta}.*
water NOM {evaporation-do-PST/evaporation ACC do-PST}
'The water evaporated.'

The unergative VN *UNDOO* 'exercise' is felicitously marked with the accusative case, but the unaccusative VN *ZYOOHATU* 'evaporate' resists accusative marking. Note that both types of VNs can be separated by a focus particle, e.g. *{UNDOO/ZYOOHATU} mo suru* [{exercise/evaporate} also do]. The researchers noted above generally account for inability of unaccusative VNs to appear as an accusative-marked object by appeal to Burzio's generalization (Burzio 1986), which says that only a predicate taking an external argument can assign accusative case. In (27b), the VN *ZYOOHATU* 'evaporate' cannot be marked with accusative case, because it does not take an external argument. Some other examples of unaccusative VNs that appear incompatible with accusative marking are *HASSEI* 'occur', *SIBOO* 'pass away', *TANZYOO* 'be born', and *ITTEN* 'change suddenly'. On the other hand, Uchida and Nakayama (1993) claim that the accusative marking is allowed only with those VNs that denote activity or accomplishment eventualities; the VNs representing state or achievement eventualities are excluded because they lack denotation of ongoing activities. The notion of ongoing activities would be remotely linked to the subject's controllability and hence to the notion of external arguments.

While the implementation of Burzio's generalization in the accusative marking of VNs appears to capture the intuitions shared by many native speakers in the middle-age or senior-age groups, Kageyama (1993) observes that it tends to lose its effect for younger speakers. Internet search, in fact, brings out a large number of instances where unaccusative VNs including *HASSEI* 'occur', *SIBOO* 'pass away', and others given above are marked with the accusative in front of the verb *suru*. It is not immediately clear whether this is an ongoing language change or merely a reflection of grammatical errors.

Another notable feature of VNs has to do with the so-called "causative-inchoative alteration". Some VN-*suru* forms such as *IDOO-suru* 'move', *HASON-suru* 'damage', *KAKETU suru* 'approve', and *SYUURYOO-suru* 'finish' may function either as intransitive or transitive predicates.

(28) a. *Kaigi*   ***ga*** *SYUURYOO-si-ta*.   (Intransitive)
     meeting   NOM   finish-do-PST
     'The meeting has ended.'

    b. *Kaigi*   ***o***   *SYUURYOO-si-ta*.   (Transitive)
     meeting   ACC   finish-do-PST
     '(They) have ended the meeting.'

Similar alternations obtain with foreign VNs as well: *uriage ga APPU-suru* [sales NOM go.up-do] 'Sales increase' vs. *uriage o APPU-suru* [sales ACC increase] 'increase sales'. While there is a huge literature discussing the causative-inchoative alternations of native verbs like *kow-as-u* 'break' (tr.) and *kow-are-ru* 'break' (intr.) (Jacobsen 1991), there is virtually no detailed study on the causative-inchoative alternation of VNs, chiefly because this lexical category does not show overt affixes signaling their transitive/intransitive status.

Finally, the VN *o suru* construction displays the 'frozen phenomenon'; when the *o*-marked VN contains no argument in its domain, it becomes insensitive to a syntactic operation, such as passivization, even though it is marked with accusative case.

(29) a. *Taroo ga Hanako to*   *SANPO o*   *si-ta*.
     Taro   NOM Hanako with stroll   ACC do-PST
     'Taro took a stroll with Hanako.'

    b. **SANPO ga*   *Taroo niyotte Hanako to*   *s-are-ta*.
     stroll   NOM Taro   by   Hanako with do-PASS PST
     'A stroll was taken with Hanako by Taro.'

This does not mean that no instances of accusative objects occurring with *suru* can be displaced. Passivization is often possible when *suru* describes an agentive action, taking an ordinary noun as its object, as exemplified by (30) (cf. Hirao 1995).

(30) a. *Taroo ga Hanako ni*   *okurimono o*   *si-ta*.
     Taro   NOM Hanako DAT gift   ACC do-PST
     'Taro gave a gift to Hanako.'

    b. *Okurimono ga*   *Taroo niyotte Hanako ni*   *s-are-ta*.
     gift   NOM Taro   by   Hanako DAT do-PASS PST
     'A gift was given to Hanako by Taroo.'

Kageyama (1991) accounts for the 'frozen phenomenon' observed for the VN *o suru* construction by means of so-called Abstract Incorporation (cf. Baker 1988), which is a type of morpho-syntactic word-formation of a VN and *suru*. Under this account, since the word-formation is in essence lexical, it is not possible to move either component of the lexical item in syntax, hence providing the reason why VN-*o suru* forms exhibit the 'frozen phenomenon' (cf. Miyamoto 1999).

3.3 Other light verbs

The most typical light verb used in the light verb constructions is *suru* 'do' (and its suppletive forms of potential *dekiru* 'can do' and honorific *nasaru* 'do'). However, Matsumoto (1996) contends that some other verbs like *hazimeru* 'begin' and *kokoromiru* 'try' also behave as light verbs.

(31) a. *Karera wa Tookyoo e bussi no YUSOO o hazime-ta.*
    they TOP Tokyo to goods GEN transport ACC begin-PST
    'They began to transport the goods to Tokyo.'

    b. *Karera wa [Tookyoo e no bussi no YUSOO] o hazime-ta.*
       they TOP Tokyo to GEN goods GEN transport ACC begin-PST
       'They began to transport the goods to Tokyo.'

This construction shows a number of distributions that parallel those found in the light verb constructions, including the argument transfer effect, as seen in (31) (see section 4). Nevertheless, unlike the light verb *suru*, verbs like *hazimeru* cannot be directly combined with VNs to give rise to compound forms of "VN-Verb" like *YUSOO-*hazimeru*.

    Another difference has to do with the so-called double accusative constraint. So far we have concentrated on two types of constructions involving VNs and *suru*: (i) "NP *no* VN *o suru*" and (ii) "NP *o* VN-*suru*", as exemplified by (24a, b), repeated below as (32a) and (32b).

(32) a. "NP *no* VN *o suru*" pattern
    Gakusei ga [kyoositu no SOOZI] o si-ta.
    student NOM classroom GEN cleaning ACC do-PST
    lit. 'The students did the cleaning of the classroom.'

    b. "NP *o* VN-*suru*" pattern
    Gakusei ga kyoositu o SOOZI-si-ta.
    student NOM classroom ACC cleaning-do-PST
    'The students cleaned the classroom.'

c. "NP o VN o *suru*" pattern
   ??*Gakusei ga kyoositu o SOOZI o si-ta.*
   student NOM classroom ACC cleaning ACC do-PST
   'The students cleaned the classroom.'

Actually, however, there is the third pattern shown in (32c), where both VN and thematic object are marked in the accusative. This pattern is generally considered to be ungrammatical due to the so-called "double accusative constraint" or "double *o* constraint" (Harada 1973). Examples like (32c), particularly in written language, would be judged nearly unacceptable. Nonetheless, Kageyama (1993) points out that the double accusative pattern is amply attested in the statements of ministers, members of the Diet, newscasters, and others speaking in a dignified style. He also shows that the effect of the double accusative constraint is greatly weakened if the VN carries a focus particle (instead of an accusative marker) and that this constraint is not applicable if the sentence is put in passive, where an accusative object is turned into a nominative subject. Observe the examples in (33) taken from Kageyama (1993: 317).

(33) a. *setubi o donoyooni KATUYOO o si-te ik-u ka*
     equipment ACC how utilize ACC do-GER go-PRS Q
     'how the equipment is made best use of''

   b. *Keisatu wa soori-daizin o TAIHO sae si kanenai.*
      police TOP prime-minister ACC arrest even do is.likely
      'It is likely that the police even arrests the Prime Minister.'

   c. *Zyookyaku no uti, zyuukyuu-nin ga KYUUZYO o s-are-masi-ta.*
      passenger GEN among nineteen-people NOM rescue ACC do-PASS POL PST
      'Of the passengers, nineteen were rescued.'

Now, the fact is that *hazimeru* 'begin' and other verbs that Matsumoto (1996) claimed to behave like light verbs do not fit into any of the three constructions in (33). It remains to be seen what kind of lexical property distinguishes the verbs pointed out by Matsumoto (1996) from the genuine light verb *suru*.

4 Argument realization in light verb constructions

In the previous section, we have seen some of the characteric properties of VN-*suru* and VN *o suru* constructions. There are certain characteristics which are unique to the VN *o suru* construction, and in this section our discussion focuses on the peculiar properties of VN *o suru* forms.

4.1 Patterns of argument realization

There are two notable properties of argument realization in VN *o suru* constructions. In the first place, the valency and argument types of VN *o suru* forms squarely match with those of their corresponding VNs, as exemplified in (34).

(34) a. *Taroo ga* SANPO *o si-ta.*
   Taro NOM stroll ACC do-PST
   'Taro took a stroll.'

   b. *Taroo ga kankoku ni* RYOKOO *o si-ta.*
   Taro NOM Korea to travel ACC do-PST
   'Taro traveled to Korea.'

   c. *Taroo ga murabito ni ookami ga kuru to* KEIKOKU *o si-ta.*
   Taro NOM villager DAT wolf NOM come COMP warning ACC do-PST
   'Taro warned the villagers that the wolf would come.'

In (34a), the VN SANPO 'stroll' has the argument structure <Agent>, and thus, the agent nominal appears. In (34b), RYOKOO 'travel' has the argument structure <Agent, Goal>, so that the sentence takes agent and goal arguments. In (34c), KEIKOKU 'warning' has <Agent, Goal, Theme>, so the clause has the three arguments of agent, goal, and theme. What is more, as discussed in section 2.2, VNs can function as predicates without *suru* if combined with *–sai* 'time'. These facts imply that *suru* combined with VNs is thematically non-significant or totally empty.

Secondly, the arguments of a VN are often realized not in the lower phrasal domain but in the upper clausal domain, without respecting 'locality', i.e. 'the argument promotion effect' is observed. In (35), RYOKOO 'travel' has <Agent, Goal>, and thus, the seeming arguments of RYOKOO 'travel' (i.e., Agent and Goal) are realized as those of the clause.

(35) *Taroo ga Tookyoo ni* RYOKOO *o si-ta.*
   Taro NOM Tokyo to travel ACC do-PST
   'Taro traveled to Tokyo.'

Although these properties of VN *o suru* construction were originally brought to linguists' attention in the early 1970s (Martin 1975), serious linguistic inquiry started with Grimshaw and Mester's (1988) seminal work with their postulation of the so-called Argument Transfer Hypothesis. Subsequent studies on the VN *o suru* construction can be divided into two groups. One group of research subscribes to the so-called 'light *suru* hypothesis', which is the view that *suru* functions as a light

verb, and the other the so-called 'heavy *suru* hypothesis', which is the view that *suru* is an ordinary heavy (i.e. not light) verb.

4.2 The light *suru* hypothesis

Grimshaw and Mester (1988) attempt to account for the fact that a VN heading an accusative noun phrase can license arguments outside its own nominal domain, as exemplified in (36).

(36) **Taroo ga   murabito ni** [ookami ga   kuru to   no KEIKOKU] o   si-ta.
     Taro   NOM villagers to wolf   NOM come COMP GEN warning ACC do-PRS
     'Taro warned the villagers that the wolf would come.'

They postulate a construction-specific mechanism termed 'Argument Transfer', by virtue of which some arguments of a VN are transferred into the argument structure of *suru* as illustrated in (37).

(37) a. KEIKOKU   (Agent, Goal, Theme)

   b. *suru* (   ) <acc>

   c. KEIKOKU   (Theme) + *suru* (Agent, Goal) <acc>
       (Grimshaw and Mester, 1988: 212)

In their analysis, the Agent and Goal arguments out of the three arguments associated with KEIKOKU 'warning', are transferred to *suru*, which is assumed to be a mere accusative case assigner. As is obvious from (37), this transferring process takes place at the lexical level, while these two arguments are θ-marked by *suru* at the syntactic level. This accounts for licensing and θ-marking of the clausal phrases (i.e., the nominative Agent phrase *Taroo* and the *ni-* ('to') marked Goal phrase).

The peculiarity of light verb constructions leads to a debate over the nature of *suru*, which occurs with accusative VN. The argument structure of a clause and that of a VN seem to correspond squarely, and this correspondence between the argument structure of VNs and the types of arguments allowed in the clause is the major reason why many researchers regard *suru* as a light verb. However, even among studies advocating the light *suru* hypothesis, there are two different views as to its 'weight'. One is that *suru* is thematically totally empty, and the other is that it is partially empty in the sense that *suru* is associated with an external argument.

While Grimshaw and Mester (1988) represent the former view, the later view is pursued by Isoda (1991), for instance. Isoda (1991) claims that there are two types of *suru*: a heavy *suru* and light *suru*. As for the light *suru*, he argues that, as shown in (38), its external argument and that of VN are fused to form a bi-clausal a(rgument)-structure (see also Matsumoto 1996; Sato 1993; Yokota 2005).

(38) a. *John wa murabito ni ookami ga kuru to KEIKOKU o si-ta.*
   John TOP villagers to wolf NOM come COMP warning ACC do-PST
   'John warned the villagers that a wolf would come.'

   b. warning-do <agent <agent goal theme>>
              | | | |
             SBJ   OBJ COMP

This (partially) fused argument structure is then mapped into the so-called Grammatical Function, realizing the fused Agent as a SBJ, and the Goal and Theme of the VN as an OBJ and a COMP, accordingly. In other words, while the light *suru* participates in the fusion (i.e., binding relationship) for its external argument, its internal arguments are supplied solely by the VN.

In regard to the 'argument promotion effect', many different mechanisms have been proposed: lexical accounts (e.g., Grimshaw and Mester 1988; Isoda 1991; Sells 1990), morpho-syntactic accounts (e.g., Kageyama 1991) and syntactic accounts (e.g., Ishii 2009; Matsumoto 1996; Nakajima 2008; Yamamoto 1992). The lexical approach is straightforward in that the lexical formation of a VN and *suru* constitutes a single predicate to license and θ-mark its clausal arguments (as in Isoda 1991). Kageyama's (1991) morpho-syntactic approach positing Abstract Incorporation is also straightforward. Under his analysis, due to the so-called Government Transparency Corollary (Baker 1988), any internal arguments, such as a Goal, are placed under the verb projection. As a result, it becomes possible to treat any internal arguments of a VN as if they were those of *suru*.

There are also various syntactic approaches: e.g. A(-rgument) Percolation (Sato 1993), Late Lexical Insertion (Yamamoto 1992), and LF Incorporation (Ishii 2009; Saito and Hoshi 2000). Saito and Hoshi (2000) claim that cyclic discharge of thematic roles eliminates any ungrammatical forms of VN *o suru*. In their analysis, the VN is raised to the position of *suru* in LF and assigns θ-roles in accordance with the Thematic Role Hierarchy. Under their analysis, the sentence in (39) is ruled out because the thematically lower argument "that the wolf was coming" (i.e. Theme) is realized in the clausal domain while the thematically higher argument "to the villagers" (i.e. Goal) stays in the nominal domain, not respecting the Thematic Role Hierarchy.

(39) **Taroo ga **ookami ga kuru to** [**murabito e** no KEIKOKU] o si-ta.*
   Taro NOM wolf NOM come COMP villagers to GEN warning ACC do-PST
   'Taro warned the villagers that the wolf was coming.'

At the same time, their analysis can account for the 'argument promotion effect' in (40) without Argument Transfer.

(40) *Mary ga John {e/ni} [$_{NP}$ toti no ZYOOTO] o si-ta.*
    Mary NOM John {toward/to} land GEN giving ACC do-PST
    'Mary gave a piece of land to John.' (Saito and Hoshi 2000: 274)

In the VN *zyooto* 'giving' first discharges its Theme role (i.e., *toti* 'land') *in situ* (i.e., within the NP), and then, the Goal role to *John* and subsequently discharging the Agent role (to *Mary*) at the highest structural position. For this approach to work, they need only adhere to the Minimalist Program's assumption that thematic roles are appropriately assigned not at D-structure (and S-structure) but in LF. This analysis eliminates a number of highly stipulative constraints that Grimshaw and Mester (1988) need to posit (see Grimshaw and Mester (1988) for the details).

4.3 The heavy *suru* hypothesis

The other group of studies, termed 'heavy *suru* hypothesis', rejects the view that *suru* functions as a light verb, but there is no consensus concerning the valency and thematic array of *suru*. In general, there are two ideas on this matter: i) that *suru* is a two-place predicate licensing Agent and Theme (or EVENT), and ii) that *suru* is a three-place predicate, additionally licensing an oblique argument, such as a Goal.

    There are two major motivations for the 'heavy *suru* hypothesis'. Firstly, Terada (1990), who is the initiator of the heavy *suru* hypothesis, observes that *suru* imposes a thematic constraint in the sense that the subject of VN *o suru* must be an Agent as in (41a), and cannot, for instance, be inanimate as in (41b).

(41) a. ***Takasi ga*** *atarashii mondai no SISA o si-ta.*
    Takashi NOM new problem GEN suggestion ACC do-PST
    'Takashi suggested a new problem.'

   b. ****Kono deetaa ga*** *atarasii mondai no SISA o si-te i-ru.*
    This data NOM new problem GEN suggestion ACC do-GER be-PRS
    'This data suggests a new problem.' (Terada, 1990: 108)

If *suru* were a (thematically totally empty) light verb, there should not be such a thematic constraint.

    Secondly, as Miyamoto (1999) argue, when a VN whose aspectual type is ACHIEVEMENT (i.e., *SYUURI* 'repair' in (42)), the VN must always assume a PROCESS reading if it heads the accusative phrase.

(42) (i) VN-*suru* form:

    Taroo wa kuruma o* **SYUURI-si-ta *ga* SYUURI-deki-nakat-ta.

    Taro TOP car ACC repair-do-PST but repair-can-NEG PST

    lit. 'Taro repaired the car but (he) could not repair (it).'

  (ii) VN *o suru* form:

    Taroo wa [kuruma no **SYUURI]** **o** **si-ta** *ga* SYUURI-deki-nakat-ta.

    Taro TOP car GEN repair ACC do-PST but repair-can-NEG PST

    lit. 'Taro repaired the car but (he) could not repair (it).'

In (42ii) is acceptable, since the VN *o suru* form has no entailment of the resulting state (i.e. 'a car having been repaired'). But since such a PROCESS reading is not possible with the contrasting VN-*suru* form, (42i) is semantically incongruent. If *suru* were to function as a light verb, the verb would be aspectually totally non-significant, and such an aspectual constraint would not exist for the VN *o suru* form.

One potential obstacle any heavy *suru* hypothesis faces up to is how to account for the so-called 'argument promotion effect'. The common strategies are to regard seemingly promoted satellites as i) arguments of *suru*, ii) the outcome of complex predicate formation, or iii) adjuncts.

Hasegawa (1991) opts for the idea that *suru* is a three-place predicate, selecting an Agent subject, EVENT object, and, optionally, an indirect Goal object, just like *tuzukeru* 'to continue' or *okonau* 'to execute', as in (43).

(43) a. *Taroo ga (Hanako ni) sore o {si-ta/tuzuke-ta/okonat-ta}.*

    Taro NOM Hanako DAT that ACC {did-PST/continue-PST/execute-PST}

    lit. 'Taro {did/continued/executed} that (to Hanako).'

  b. *Taroo ga Hanako ni kookana okurimono o {si-ta/?tuzuketa}.*

    Taro NOM Hanako DAT expensive present ACC {do-PST/continue-PST}

    lit. 'Taro {made/continued} an expensive gift to Hanako.'

Terada (1990) also opts for the view that *suru* is a three-place predicate consisting of an Agent, Theme, and, additionally, oblique argument. Under the three-place predicate hypothesis, a goal, which looks like an extra argument arising from the argument promotion effect, is taken to be an argument to the verb *suru*.

Uchida and Nakayama (1993) hold that *suru* is a two-place predicate, taking subject and object. Uchida and Nakayama (1993) account for the argument-promotion effect of PPs by 'predicate formation' and for that of CPs by merely treating them as 'adjuncts'. In (44), there is a Goal phrase (i.e., *otooto ni* 'to brother'), which cannot be licensed by the action noun *itazura* 'trick' lacking a thematic content.

(44) *John ga* **otooto ni** *itazura o si-ta.*
John NOM brother DAT trick ACC do-PST
'John played a trick on his brother.'

Uchida and Nakayama (1993: 653) claim that "the heavy *suru* is capable of assigning the goal theta role depending on the direct object, even though the direct object does not have an argument structure." Based on this idea, they view Goal PPs, such as the one in (45), as being licensed by the whole string of VN *o suru*.

(45) *John wa* **murabito ni** *[CP ookami ga kuru to] [NP KEIKOKU] o si-ta.*
John TOP villager DAT wolf NOM come COMP warning ACC do-PST
'John warned the villagers that the wolf was coming.'

As for the promotion effect of the CP (i.e., *ookami ga kuru to* 'that the wolf is coming'), Uchida and Nakayama (1993) utilize CP-movement, which adjoins a CP into the matrix VP node, as in (46).

(46) $[_{VP}$ CP$_i$ $[_{VP}$ NP (Goal) $[_{V'}$ $[_{NP}$ t_i $[_N$ KEIKOKU V]]]
(Uchida and Nakayama, 1993: 658)

According to them, this type of movement is allowed with any heavy verbs, including *suru*. In short, Uchida and Nakayama (1993) attribute the argument promotion effect involving PPs to 'predicate formation' and the one involving CPs to 'VP-adjunction'.

4.4 Thematic properties of VNs

While most studies on Japanese light verb constructions rely on the idea that the verb *suru* appearing on VN *o suru* is either a light verb or a heavy verb, Miyamoto (1999) advances an alternative view to the effect that VNs rather than the verb *suru* play a key role in VN *o suru* forms. He proposes that a VN could be either non-thematic (simple event) nominal or thematic (complex event) nominal, building on Grimshaw's (1990) work, and that VN *o suru* forms with a non-thematic VN are called 'mono-predicational', as they contain one predicate (i.e. *suru*) and those with a thematic VN are called 'bi-predicational', as they contain two predicative elements (i.e. VN and *suru*).

The two types of VN *o suru* constructions show distinct behavior, and the distinction of the thematic vs. non-thematic VNs can for instance be differentiated in two steps: firstly by modifying VN phrases either by referential or tempo-aspectual modifiers, and secondly by putting these VN phrases into temporal adjunct clause or nominal control constructions. To be concrete, first, a given VN can assume a referential modifier (i.e., *takumina* 'skillful') as in (47a), or a tempo-aspectual modifier (i.e., *nitiya* 'day and night') as in (47b).

(47) a. *Hanako no idengaku no* **takumina** *ZIKKEN*
　　　Hanako GEN genetics GEN skillful experiment
　　　'Hanako's skillful experiment of genetics'

　　b. *Hanako no idengaku no* **nitiya** *no ZIKKEN*
　　　Hanako GEN genetics GEN day.night GEN experiment
　　　'Hanako's day and night genetics experimenting'

As seen in (48), once these VN phrases are set into the temporal adjunct clause construction (e.g., *tyuu* 'during'), the grammatical difference becomes manifested.

(48) a. **Hanako no idengaku no* **takumina** *ZIKKEN-***tyuu**
　　　Hanako GEN genetics GEN skillful experiment-during
　　　lit. 'While Hanako's skillful experiment of genetics, . . .'

　　b. *Hanako no idengaku no* **nitiya** *no ZIKKEN-***tyuu**
　　　Hanako GEN genetics GEN day.night GEN experiment-during
　　　lit. 'While Hanako's day and night genetics experimenting, . . .'

Note that when a VN is combined with *tyuu*, the verb *suru* does not show up. Given the crucial fact that the test construction is compatible only with a thematic nominal (Iida 1987), the grammatical contrast assures that the VN in (48a) assumes a non-thematic reading and its isomorphemic VN in (48b) takes a thematic reading.

　　Furthermore, as seen in (49), the identical grammatical contrast is manifested when the VN phrases are set into another diagnostic tool, i.e., the nominal control construction with *tame no* ('for the sake of'), which is also compatible only with a thematic nominal.

(49) a. **[[tokkyo o toru tame] no [_{NP} Hanako no idengaku no*
　　　patent ACC obtain sake GEN Hanako GEN genetics GEN

　　　takumina *ZIKKEN]]*
　　　skillful experiment
　　　lit. 'Hanako's skillful experiment of genetics to obtain a patent'

　　b. *[[tokkyo o toru tame] no [_{NP} Hanako no idengaku no*
　　　patent ACC obtain sake GEN Hanako GEN genetics GEN

　　　nitiya *no ZIKKEN]]*
　　　day.night GEN experiment
　　　lit. 'Hanako's experimenting of genetics for day and night to obtain a patent'

The contrast in grammaticality demonstrated by these two diagnostic constructions assures that a given VN could be isomorphemically non-thematic or thematic (see Miyamoto 1999 for details).

5 A corpus study

In Miyamoto (1999) it is claimed that VN *o suru* forms can be either mono-predicational or bi-predicational, the difference being determined by the thematic property of VNs. In this section, we report the results of a small-scale corpus survey conducted from the perspective of Miyamoto (1999), with a view to checking whether there is any viable distributional difference based on mono- and bi-predicational distinction. The corpus study, among a number of findings, brings to light the fact that mono-predicational VN *o suru* forms are more common in occurrence than bi-predicational VN *o suru* forms.

5.1 Methods and analysis

The corpus used for the survey is the "Balanced Corpus of Contemporary Written Japanese (BCCWJ)", which includes 100 million words from a wide variety of written Japanese texts dating from the middle of the 19th century up to the present era. Among the Japanese verbs listed in the BCCWJ, *suru* 'do' was the second most frequent verb, and the total entry number of this verb was 380,059. The actual number of '*o suru*' string entries was 142,759. To conduct our analysis, we selected and analyzed three VN *o suru* forms: *SINPAI o suru* 'to worry', *HYOOKA o suru* 'to evaluate' and *HOOKOKU o suru* 'to report'. We chose these three forms because the frequencies of these forms were almost identical. Their frequencies were also not significantly high, so the actual data size we would have to deal with would not be voluminous, making our task easier to handle. Furthermore, to set them on equal ground, we chose and analyzed the first 100 entries for each of the VN *o suru* forms.

For the sake of simplicity, we employed the following operational criteria to differentiate between mono-predicational VN *o suru* forms (with non-thematic VNs) and bi-predicational VN *o suru* forms (with thematic VNs). Firstly, the VN *o suru* forms whose VNs had no modifiers were categorized as 'bi-predicational', as in (50).

(50) *Nanika   at-ta-ra   sugu-ni   **HOOKOKU o su-ru**   yo.*
anything is-PST COND immediately report   ACC do-PRS PRT
'If there is anything, (I will) give you my report as soon as possible.'

Secondly, the VN *o suru* forms whose VNs had modifiers were categorized as 'mono-predicational' specifically when these modifiers were not accompanied by the possessive case of *no*, as in (51).

(51) *Ragu wa Silbetto ni **turai** HOUKOKU o si-nakerebanaranakat-ta.*
Lag TOP Silvet DAT painful report   ACC do-must-PST
'Lag had to let Silvet know the bad news.'

Thirdly, we excluded the VN *o suru* forms whose VNs had possessive (*no*)-modifiers, such as in (52) due to the ambiguous status between mono- and bi-predicationality.

(52) *Byoonin   ni   **kane   no** SINPAI o   s-ase-ru   no   wa ichiban-warui.*
　　 sick.person DAT money GEN worry ACC do-CAUS PRS COMP TOP most-bad
　　 'There is nothing worse than making a sick person worry about money.'

Finally, the VN *o suru* forms with compounding VNs such as (52) were categorized as 'mono-predicational'.

(53) *Tantoo   yakuin ga   **kessan-HOOKOKU o   si-te**   syoonin o   e-masu.*
　　 in.charge director NOM closing-report   ACC do-GER approval ACC gain-POL
　　 'The director in charge will gain approval after making a report on the closing.'

To see if there was any statistical significance concerning the occurrence of the thematic and non-thematic VN *o suru* forms, we conducted series of chi square tests. The observed frequencies of the three VN *o suru* forms are summarized in Table 3.

Table 3: Frequencies of mono- vs. bi-predicational *VN o suru* forms

	SINPAI o suru	*HYOOKA o suru*	*HOOKOKU o suru*
Verbal Noun	23	32	36
Modifier + Verbal Noun	43	35	35
Compound VN	0	19	17

In Table 3, "Verbal Noun" (i.e., VN *o suru* forms without modifiers in their accusative nominal domains) refers to the frequency of bi-predicational VN *o suru* forms. Contrastively, "Modifier + Verbal Noun" and "Compound VN" refer to the frequency of mono-predicational VN *o suru* forms.

In the case of *SINPAI o suru* 'to worry' ($\chi^2(1) = 6.061$, $p < .05$) and *HYOOKA o suru* 'to evaluate' ($\chi^2(1) = 5.628$, $p < .05$), there was a statistically significant difference in the frequency of the bi- and mono-predicational VN *o suru* forms. Even in the case of *HOOKOKU o suru* 'to report', though marginal, there was a difference in frequency between the two types of VN *o suru* forms ($\chi^2(1) = 2.909$, $p < .10$). These results indicate that the bi-predicational VN *o suru* forms are less frequent in occurrence than their mono-predicational counterparts.

Another series of chi square tests was conducted to examine whether VN *o suru* forms occurred as matrix predicates or not. Table 4 shows how often the VN *o suru* forms occurred in two contrasting positions: sentence final positions and sentence non-final (embedded) positions.

Table 4: Positional differences of VN *o suru* forms

	SINPAI o suru	*HYOOKA o suru*	*HOOKOKU o suru*
Sentence Final Positions	27	40	23
Non-final Positions	71	51	55

A series of our χ^2-tests indicated that the VN *o suru* forms were less likely to function as matrix predicates. For *SINPAI o suru* 'to worry' ($\chi^2(1) = 19.755$, *p* < .001) and *HOOKOKU o suru* 'to report' ($\chi^2(1) = 13.128$, *p* < .001), there were statistically significant differences in the frequency of occurrence between those forms which appeared at the sentence final positions and those forms which appeared in other (embedded) positions, such as in relative clauses, as exemplified in (54).

(54) *Sikasi kare no baai **i-eki** de tokas-are-ru* SINPAI *o* su-ru
but he GEN case gastric-juice INST dissolve-PASS PRS worry ACC do-PRS

***hituyoo** wa nakat-ta.*
necessity TOP not.exist-PST
'But in his case he did not have to worry about (his stomach) to be dissolved by the gastric juice.'

Even though this tendency was not so obvious with *HYOOKA o suru* 'to evaluate' ($\chi^2(1) = 1.330$, *p* = .25), overall, these three VN *o suru* forms occurred less frequently sentence-finally than in other (embedded) positions ($\chi^2(1) = 28.348$, *p* < .001).

Another finding of ours was the occurrence of 'VN-compounds', such as KATI HYOOKA '(lit.) value-evaluation'. Although it is beyond the scope of this study to discuss their theoretical significances, at least the following morpho-syntactic patterns exist. While the primary element (on the right side) is functioning as a predicate, the secondary element functions either as its argument, adjectival or adverbial, as in (55).

(55) i. Arguments:
    a) KATUDOO HOOKOKU 'activity-report'
      ($\approx$ *katudoo o hookoku suru* [activity ACC report do] 'report the activity')
    b) TAIKEN HOOKOKU 'experience-report'
    c) ZIKO HOOKOKU 'accident-report'

  ii. Adjectivals:
    a) TYUUKAN HYOOKA 'mid-evaluation'
      ($\approx$ *tyuukanteki-na hyooka* [intermediate report] 'interim report')
    b) KYAKKANTEKI HYOOKA 'objective-evaluation'
    c) KOO HYOOKA 'good-evaluation'

iii. Adverbials:
a) KADAI HYOOKA 'excessive-evaluation'
   (≈ *kadai-ni hyooka suru* [excessively evaluation do] 'overestimate')
b) KASYOO HYOOKA 'too-little-evaluation'
c) HIKAKU HYOOKA 'comparative-evaluation'

Although VN-compounds were not uncommon in our corpus data, the literature of VN *o suru* is silent about how and why this type of VN should occur.

5.2 Loan words as VNs

Loanwords such as those from English have been increasing their presence in the current use of Japanese as part of its vocabulary. This tendency was also observed to a certain degree with VN *o suru* forms (e.g., *DORAIBU o suru* 'to drive' and *ASISUTO o suru* 'to assist'). The BCCWJ corpus confirms that the number of loan words used as VNs has been steadily increasing in recent years.

To see general trends in use, we gleaned the VN *o suru* formed with loanwords, and listed all of these forms as Table 5.

We made a number of findings from the data, outlined here. First, the number of loanwords in the form of VN *o suru* has steadily been increasing, many of which reflect changes in lifestyle and technology. For instance, *MEIRU o suru* 'to e-mail', shown in (56), occurs frequently due to advances in information technology.

(56) *Hoka   no   zibun no   meiru adoresu ni   **MEIRU o   su-ru**.*
      another GEN   self GEN mail address DAT mail   ACC   do-PRS
      '(I'm going to) send the e-mail to another email address of mine.'

Secondly, some of the VNs of Japanese or Chinese origin have been replaced with loanword equivalents. For instance, *ANMA o suru* 'to massage' has been replaced with *MASSAAZI o suru* 'to massage'. Also, instead of using *HANASI o suru* 'to talk', its loanword version of *TOOKU o suru* 'to talk' has been becoming more favored in use.

(57) *Utiyama   ga   mezasu zyosei-zoo   nituite **TOOKU o   su-ru**.*
      Uchiyama   NOM   aspire female-image   about   talk   ACC   do-PRS
      'Uchiyama talks about the image of (his) aspiring female.'

Thirdly, there are loanwords that have become thoroughly entrenched in Japanese. For instance, *SAIN o suru* 'to sign' and *KISU o suru* 'to kiss' have been used for over 20 years as part of the Japanese vocabulary; also the frequency of *KISU o suru* has been steadily getting higher. To be more specific, 32 out of 134 entries of loanwords

Table 5: Entries of loanwords formed with "VN *o suru*"

Total Frequency	Year	Entries				
3	1976	PR				
1	1977	knock				
3	1981	recording				
1	1986	catch ball				
5	1987	sign	massage	game	play	
5	1988	pitch	kiss	sex		
4	1989	memo	change	call	running	stitch
3	1990	billiard	gesture			
9	1991	ear ring				
8	1992	demo	speech	sketch	check	naming
10	1993	comment	engine cut	baton touch	wink	OK
8	1994	lesson	camp	dance	pass	clear kick
7	1995	date				
10	1996	service	sales	patrol	presentation	multi
11	1997	wiring	garland	*twin sting method*		
5	1998	simulation	picnic	advice		
15	1999	training	volunteer	*guts pose*	SM	
14	2000	shopping	communication	*H*		
70	2001	U turn	golf	*Arbeit*	propose	guiding
		try	click	office work	gear change	setting
		performance	long jump	restructure	*caraméliser*	
43	2002	television game	running	coordinate	running stitch	request
		test	dribble	miss	meeting	
62	2003	loca(tion)hun(ting)	guard	*enquête*	pose	announce
		lock	clear	workshop	cross stitch	attack
		collection	baton touch	reaction	warm up run	quilting
		wrestling	dial	wrap	*pachinko*	rehearsal
74	2004	dash	present	jogging	swing	hitchhike
		race	free talk	check in	support	self talk
		design	reform	self conditioning		
135	2005	mail	closing	custom	clean up	necktie
		set up	American football	talk	image	sunglass
		DJ	CM	LAN	MRI	PM
		professional wrestling		BBQ	TOB	FAX
2	2006	jump				
29	2008	cut	approach	tennis	deep kiss	receive
		message	mask			

recorded in 2005 were the *KISU o suru* form, illustrating that this expression has firmly taken its root in Japanese.

Finally, the number of loanwords written in the Roman alphabet has been increasing. Our data shows that this trend has been especially noticeable since 2000. To be more specific, out of 96 loanword entries that were rendered in the Roman alphabet, 85 were made in the 2000s. *DJ o suru* (58a) and *TOB o sita* (58b) are a couple of these forms.

(58) a. *Wakai DJ-tati ga   atsumat-te DJ o   su-ru.*
      young DJ-PL   NOM gather-GER DJ ACC do-PRS
      'Gathering together, young DJs play music.'

   b. *Huzi terebi wa TOB   o   si-ta.*
      Fiji TV   TOP take-over-bid ACC do-PST
      'Fuji TV station made a takeover bid.'

Excluding those with strong sexual connotation, Table 5 lists all of the entries of VN *o suru* formed with loanwords that we gleaned from the Corpus. (Note that VN-*suru* forms made with loanwords were not part of our research scope.) 'Entries' refers to the loanwords associated with the VN *o suru* form; for instance, 'PR' refers to the entry of '*PR o suru*'. These entries are placed under the 'Year', which refers to their appearance in the BCCWJ for the first time. Note that in the Table, each entry appears only once even if it may have appeared in multiple years or even every year. 'Total Frequency' refers to the number of appearance for given entries; for instance, in the year of 1976, *PR o suru* appeared three times. In the Table, non-English forms are italicized.

5.3 Results

The small-scale corpus study brings to light a few new findings. First, we have discovered that mono-predicational VN *o suru* forms are more common in occurrence than bi-predicational VN *o suru* forms. Additionally, the survey reveals that regardless of VN *o suru* types, these forms were less likely to occur as matrix predicates. Rather, they were embedded, predominantly in relative clauses, and there were a fair amount of entries in which VNs themselves were compounds. Lastly, the number of entries has been steadily increasing in recent years. Admittedly, the survey was very limited in scope and size, and we also have to leave for further study the question of why the VN *o suru* forms should have the distributions noted above. In any event, what this section illustrates is that corpus studies are often instrumental in shedding light on some facts that have not drawn much attention in theoretical studies.

6 Conclusion and future research perspectives

In this chapter, we have discussed the light verb constructions which comprise VNs and the light verb *suru*, and numerated a number of peculiar characteristics associated with the light verb constructions. We have seen some notable features of the verb *suru* and VNs in Sections 2 and 3. Section 4 has been devoted to the discussion of the VN *o suru* construction, and illustrated how the peculiar behaviors of this construction are accounted for by the light *suru* hypothesis and heavy *suru* hypothesis.

Section 5 is a sample corpus study to fathom the nature of the VN *o suru* construction, and we have discussed a number of findings.

There are many issues that are worth further exploration. While the discussion in this chapter places a particular emphasis on VN *o suru* forms, the compound variant of the VN-*suru* form has many interesting properties that merit further intensive investigations. Since VN *o suru* and VN-*suru* carry the same meaning, and since both function as complex predicates in the same way despite the difference in morphological form, an issue arises as to how the two forms are related. Moreover, at first blush, the compound form VN-*suru* seems to have a status of 'word', but this is not so obvious, because VN-*suru* must somehow have a periphrastic structure when it is nominalized. This fact also raises an interesting theoretical question of why this must be the case.

Many theoretical studies have agreed that VNs should be equipped with their own argument structure, but no consensus has been reached in the status of the verb *suru*. It is beyond doubt that many intricate behaviors observed in the light verb constructions are derived from the special character of VNs. Nevertheless, the precise role that *suru* plays in the light verb construction is controversial. In particular, in the presence of intriguing empirical data, it is not easy to determine the exact nature of the verb *suru* combined with VNs, as well as to provide a principled account for how the complex forms VN *o suru* and VN-*suru* forms allow arguments to appear within and outside VNs. A definitive theory that can account for the intriguing behaviors of the light verb construction with full generality is yet to come. On top of this fact, *suru* has many other intricate uses, so it might be worthwhile to pursue the question of what property gives rise to various uses of *suru*. There is also a question as to what kind of lexical property allows verbs other than *suru* (e.g. verbs like *hazimeru* 'begin' and *kokoromiru* 'try') to function as light verbs.

It is worthy of remark here that many languages have light verb constructions comparable to the Japanese light verb construction, and that they are intensively studied from many different perspectives (see e.g., Alba-Salas 2004; Bak 2011; Butt and Geuder 2001; Karimi-Doostan 1997; Kearns 2002). Since the light verb constructions in other languages have properties that are not observed in the Japanese light verb constructions, the comparative or cross-linguistic studies of light verb constructions will be beneficial as the future research, and surely bring us further insight into the nature of the light verb constructions.

Acknowledgements

References

Akita, Kimi and Natsuko Tsujimura. this volume. Chapter 4: Mimetics. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Alba-Salas, Josep. 2004. *Fare* light verb constructions and Italian causatives: Understanding the differences. *Rivista di Linguistica* 16(2). 283–323.

Bak, Jaehee. 2011. *The light verb construction in Korean*. Toronto: University of Toronto dissertation.

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Botha, Rudolph. 1984. *Morphological mechanisms: Lexicalist analyses of synthetic compounding*. Oxford: Pergamon Press.

Burzio, Luigi. 1986. *Italian syntax*. Dordrecht: Reidel.

Butt, Miriam and Wilhelm Geuder. 2001. On the (semi) lexical status of light verbs. In Norbert Corver and Henk van Riemsdijk (eds.), *Semi lexical categories: On the content of function words and the function of content words*, 323–370. Berlin: Mouton de Gruyter.

Dubinsky, Stanley. 1989. Compound *suru* verbs and evidence for unaccusativity in Japanese. In Caroline Wiltshire, Randolph Graczyk and Bradley Music (eds.), *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society Part One: The General Session*, 98–111.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Grimshaw, Jane and Armin Mester. 1988. Light verbs and θ-marking. *Linguistic Inquiry* 19. 205–232.

Harada, S.I. 1973. Counter equi NP deletion. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics* 7. 113–147. University of Tokyo.

Hasegawa, Nobuko. 1991. On head movement in Japanese: The case of verbal nouns. *Proceedings of SLS* 6. 8–33.

Hirao, Tokuko. 1995. VN ga suru to VN-suru to VN o suru [VN ga suru and VN-suru and VN o suru]. In Miyajima Tatsuo and Nitta Yoshio (eds.), *Nihongo ruigi hyōgen no shosō* [Aspects of synonymous expressions in Japanese], 89–98. Tokyo: Kurosio Publishers.

Iida, Masayo. 1987. Case-assignment by nominals in Japanese. In Masayo Iida, Stephen Wechsler and Draga Zec (eds.), *Working papers in grammatical theory and discourse structure*, 93–138. Stanford: CSLI Publications.

Ishii, Toru. 2009. On PF-LF mismatch in the Japanese light verb construction. *Language and Linguistics* 10. 629–667.

Isoda, Michio. 1991. The light verb construction in Japanese. In Lisa M. Dobin, Lynn Nichols and Rosa M. Rodriguez (eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistic Society 1991 Part One: The General Session*, 261–275.

Jacobsen, Wesley M. 1991. *The transitive structure of events in Japanese*. Tokyo: Kurosio Publishers.

Jespersen, Otto. 1966. *A modern English grammar on historical principles*. London: George Allen and Unwin.

Kageyama, Taro. 1980. *Goi no kōzō* [Lexical structures]. Tokyo: Shōhakusha.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1991. Light verb constructions and the syntax-morphology interface. In Heizo Nakajima (ed.), *Current English linguistics in Japan*, 169–203. Berlin: Mouton de Gruyter.

Kageyama, Taro. 1992. Dummy *su-* and lexical *su-*: A reply to Miyara. *Gengo Kenkyu* 102. 165–172.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1999. Word formation. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 297–325. Malden: Blackwell.

Kageyama, Taro. 2004. Keidōshikōbun to shiteno aoi me o shiteiru kōbun. [“Having blue eyes” sentences as a light verb construction]. *Nihongo bunpō* [Japanese grammar] 4. 22–37.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Karimi-Doostan, Gholamhossein. 1997. *Light verb constructions in Persian*. Essex, UK: University of Essex dissertation.

Kearns, Kate. 2002. Light verbs in English. Ms. MIT.

Kishimoto, Hideki. 2006. Japanese syntactic nominalization and VP-internal syntax. *Lingua* 116. 771–810.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kobayashi, Hideki, Kiyo Yamashita and Taro Kageyama. this volume. Chapter 3: Sino-Japanese words. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Miyamoto, Tadao. 1999. *The light verb construction in Japanese: The role of the verbal noun*. Amsterdam: John Benjamins.

Miyagawa, Shigeru. 1987. Lexical categories in Japanese. *Lingua* 73. 29–51.

Miyagawa, Shigeru. 1989. Light verbs and the ergative hypothesis. *Linguistic Inquiry* 20. 659–668.

Nakajima, Takashi. 2008. Loan word syntax: A case in the light verb construction. *Toronto Working Papers in Linguistics* 28. 259–280.

Saito, Mamoru and Hiroto Hoshi. 2000. The Japanese light verb construction and the Minimalist Programs. In Roger Martin, David Michael and Juan Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 261–295. Cambridge, MA: MIT Press.

Sato, Yutaka. 1993. Argument transfer: All or nothing. *Gengo Kenkyu* 104. 92–127.

Sells, Peter. 1990. More on light verbs and θ-marking. Ms. Stanford University.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar. *Language* 64. 451–484.

Sugimura, Mina. 2012. Root vs. n: A study of Japanese light verb construction and its implications for nominal architecture. In Enrico Boone, Martin Kohlberger and Maartje Schulpen (eds.), *Proceedings of 17th conference of the student organization of linguistics in Europe*, 289–298.

Tanomura, Tadaharu. 1988. Heya o sōji suru to heya no sōji o suru [Heya o sōji suru and heya no sōji o suru]. *Nihongogaku* 7(11). 70–80.

Terada, Michiko. 1990. *Incorporation and argument structure in Japanese*. Amherst, MA: University of Massachusetts dissertation.

Tsujimura, Natsuko. 1990. Ergativity of nouns and case assignment. *Linguistic Inquiry* 21. 277–287.

Uchida, Yoshiko and Mineharu Nakayama. 1993. Japanese verbal noun constructions. *Linguistics* 31. 623–666.

Yamamoto, Hiromi. 1992. On the so-called 'light' verb suru in Japanese. *University of Washington Working Papers in Linguistics* 10. 324–339.

Yokota, Kenji. 2005. The structure and meaning of Japanese light verbs. *Language Sciences* 27. 247–280.

Corpus

BCCWJ (Balance Corpus of Contemporary Written Japanese). 2011. Tokyo: National Institute for Japanese Language and Linguistics. http://www.kotonoha.gr.jp/shonagon/

Koichi Takezawa
13 Inflection

1 Introduction

The main purpose of this chapter is to elucidate the characteristics of inflection in Japanese with a focus on its syntactic aspects. The term "inflection" broadly refers to the morphological modification of the form of a word according to the syntactic environment in which it appears. In many languages, not only verbs but also nominal elements (nouns, pronouns, determiners) change their forms depending on their grammatical function, and it is customary to call the inflection of nominals "declension". In Japanese, however, where word classes are traditionally divided into two major groups, *yōgen* (inflecting words, i.e. verbs and adjectivals) and *taigen* (non-inflecting words, i.e. nouns), according to the presence or absence of inflection, only verbs and adjectivals are considered to have inflection or "conjugation" (*katsuyō*). The case particles appended to nouns or noun phrases in Japanese are radically distinct from the declensional suffixes attached to nouns in Latin and other inflectional languages, because they have syntactic independence as clitics or *fuzoku-go* ("non-independent words"), as discussed in Chapter 14 (Kageyama, this volume). Consequently, the present chapter addresses itself exclusively to the conjugation of verbs and adjectivals, with a view to clarifying how inflectional phenomena in Japanese can be characterized in syntactic terms. In particular, the correspondences and discrepancies between configurational syntactic structure and a variety of inflectional endings are discussed with special reference to phenomena that are said to be unique to Japanese.

This chapter is organized as follows. Section 2 looks into verbal inflection, which has a long history of study in the domestic tradition of Japanese grammar. After briefly reviewing how verbal inflection phenomena in Modern Japanese were analyzed in the pre-generative (traditional and American structuralist) contexts, it is shown how they can be accommodated syntactically from a generative viewpoint, based on the morphological classification made by Bloch (1946a, b). Section 3 turns to adjectives and examines their categorial and inflectional properties. Japanese adjectives are said to be unique in that there are two conjugation types; the first group is traditionally considered to have their own inflectional endings like verbs, and the second (called "Adjectival Noun") takes the inflecting copula, as do the adjectives in English and many other languages. Two opposing analyses for the treatment of the two types have been put forward: one distinguishing the two types syntactically, and the other taking them as one and the same syntactic category (see Chapter 2 [Kishimoto and Uehara, this volume] for an overview of the lexical treatments of A and AN). I will introduce Nishiyama's (1999) analysis in the Distributed Morphology

framework, where it is argued that there is no need to posit two distinct categories. Section 4 extends the discussion in Sections 2 and 3 to non-finite inflections of the verbs and adjectivals used as "secondary predicates" in resultative and depictive constructions. In Japanese, adjectival resultative phrases appear with the *ni* or *ku* endings, and depictive phrases with the V-*te* and X-*de* forms. It is demonstrated that the *ni/ku* endings in resultative predicates are allomorphic variants of the infinitive copula licensed within the domain of the VP. Also, the contrast between the syntactic and categorical status of X-*de* type depictive predicates and the V-*te* type depictives is discussed. Section 5 will summarize the discussion in the chapter and discuss remaining issues. Overall, it will be argued that Japanese conjugation systems of verbs, adjectivals, and the copula should be treated in terms of the hierarchical configuration in syntax consisting of Tense, Modal, Complementizer and other functional categories.

2 Verbal inflection

We will start with verbal inflection. Japanese verbs lack inflections triggered by grammatical agreement with subjects in person, number, or gender; instead, they are accompanied with a variety of suffixal endings that are related to finiteness of sentences, such as tense, aspect, and mood. In this section, we will observe the kinds of inflectional endings on verbs with particular reference to the celebrated classification by structural linguist Bernard Bloch (1946a, b), and discuss how they could be improved from the viewpoint of generative syntax. Before examining Bloch's classification, however, it is in order to give a brief overview of the treatments of verbal inflection in the domestic tradition of Japanese grammar, where the taxonomy of word classes (lexical categories) and classification of inflectional endings are two major issues that have been continually debated since the 18th century.

2.1 Traditional Japanese grammar

The study of conjugation, called *katsuyō-ron*, has a long history, dating back to *Kokugaku* ('Japan studies') School in the late Edo period (18th century), and constituted a central part of the philologists' research into Classical Japanese, where words were segmented by using *kana* syllabary. To name only a few distinguished grammarians of the day, MOTOORI Norinaga (本居宣長, 1730–1801), FUJITANI Nariakira (富士谷成章, 1738–1779), SUZUKI Akira (鈴木朖, 1764–1837), MOTOORI Haruniwa (本居春庭, 1763–1828), and TŌJŌ Gimon (東条義門, 1786–1843) are commonly assumed to have laid the foundation of conjugation studies in Japanese

philology. The study of inflection in Modern Japanese began at the beginning of the 20th century, when the same methodology as used in the analysis of Classical Japanese was applied to the modern language. The conjugation patterns proposed by early grammarians, notably Hashimoto (1948), still form a basis of the grammatical instruction in Japanese schools, despite criticism over its direct application of Classical Japanese grammar to the modern system. A typical schoolroom conjugation paradigm is shown in Table 1. In perusing the table, special attention should be paid to the conjugated forms of verbs and their syntactic environments in which they occur. For illustrative purposes (explained below), the conjugated forms are represented in small capitals with a dot indicating the boundary of moras.

Table 1: Verb conjugation chart in school grammar

	go-dan (penta-grade) conjugation	*ichi-dan* (mono-grade) conjugation	*henkaku* (irregular) conjugation		**syntactic environments**
citation form	*kaku* 'write'	*okiru* 'rise'	*kuru* 'come'	*suru* 'do'	
mizen (irrealis)	KA.KA KA.KO	O.KI	KO	SI	*nai* (negative) *o/yoo* (hortative)
ren'yō (adverbal)	KA.KI KA.I	O.KI	KI	SI	*masu* (polite) *ta* (past)
shūshi (conclusive)	KA.KU	O.KI.RU	KU.RU	SU.RU	#
rentai (attributive)	KA.KU	O.KI.RU	KU.RU	SU.RU	noun
katei (hypothetical)	KA.KE	O.KI.RE	KU.RE	SU.RE	*ba* ('if')
meirei (imperative)	KA.KE	O.KI.RO	KO.I	SI.RO	#

= sentence boundary

A unique feature of the traditional approach that separates it from modern linguistic approaches resides in its reliance on the *kana* syllabary system in the description of the inflectional paradigm. In Table 1, each mora unit separated by a dot, such as KA, KU, KE, I, RU, etc., is written with a separate *kana* letter. Because of this, if a morpheme division lies inside a single mora, typically between the C and V that constitutes a mora, it is impossible to extract and represent that morpheme in this system. For example, in the conclusive form of 'write', *KA.KU*, the linguistic segmentation will separate the first three sounds /kak/ as a morpheme, i.e. the stem of the verb, from the last vowel /u/ as an inflectional ending, as we will see shortly. Heavy reliance on *kana* writing thus blinded the analysts in the earliest days from seeing the correct forms of verb stems and conjugational endings.

Table 1 shows two major conjugational types, called *go-dan* (penta-grade) verbs and *ichi-dan* (mono-grade) verbs, plus two irregular verbs, *ku(ru)* 'come' and *su(ru)*

'do', which change forms according to six inflectional categories: irrealis, adverbial, conclusive, adnominal, hypothetical, and imperative. Apart from the morphological limits of segmentation based on *kana* writing, the classification of inflections into these six categories is also considered problematic. As can be seen from the labeling of the six inflectional classes, the criteria for categorization are arbitrary and inconsistent, and it has been repeatedly pointed out that they are one of the causes for confusion in the study of Modern Japanese verb conjugation. The reader is referred to Shibatani (1990) and Kinsui (1997) for critical appraisal of the traditional *katsuyō* analyses.

On the other hand, several traditional grammarians, who were aware of the limits of *kana* syllabary transcription, began to depart from the *kana*-based approach as early as the beginning of the 20th century. An outstanding scholar of the day was Yamada (1908), who pointed out the usefulness of alphabetical transcription for the study of conjugation, and Sakuma (1936) attempted a phoneme-based analysis. Among the pioneers who attempted at a more systematic description of Modern Japanese conjugation, Bernard Bloch was the first to present a principled system of inflectional alternations by rigorously applying the methodology of American structural linguistics. His analysis not only brought a consistent morphological division of stems and inflectional endings but also exerted a deep influence on subsequent studies on the morphosyntax of conjugation, notably Mikami (1963) and Teramura (1984). In the next subsection, we give a brief review of Bloch's two seminal papers (1946a, b).

2.2 Bernard Bloch

Following the methodology of morpheme segmentation in structural linguistics, Bloch (1946a) first defined two major classes of verbs in terms of stem-final sounds,

Table 2: Bloch's (1946a) verb conjugation paradigm

	Vowel-final stem 'eat'	Consonant-final stem 'read'
Non-past indicative	tabe ru	yom u
Non-past presumptive	tabe yoo	yom oo
Imperative	tabe ro	yom e
Provisional	tabe reba	yom eba
Infinitive	tabe Ø	yom i
Past indicative	tabe ta	yon da
Past presumptive	tabe taroo	yon daroo
Conditional	tabe tara	yon dara
Alternative	tabe tari	yon dari
Gerund	tabe te	yon de

namely consonant-ending verbs and vowel-ending verbs, which correspond respectively to *go-dan* (penta-grade) and *ichi-dan* (mono-grade) verbs in the traditional classification. Consonant-ending verbs are like *kak-* 'write', *yom-* 'read', and *sagas-* 'look for', where the stems end in consonants, and vowel-ending verbs are like *tabe-* 'eat', *mi-* 'see', and *sute-* 'throw away', where the stems end in vowels. On the basis of this division, Bloch identified ten inflectional forms, as shown in Table 2, where a verb stem and a conjugational ending are separated by a hyphen.

In Bloch (1946b), the ten inflectional endings he identified morphophonemically are classified into subtypes according to syntactic criteria; they are divided into two upper-groups according to their clause types, and five subgroups according to mood, as illustrated in Table 3:

Table 3: Bloch's (1946b) inflectional suffixes

clause types	mood	tense	Inflectional suffixes
final clauses	Indicative	Non-past	*ru, u*
		Past	*ta, da*
	Presumptive	Non-past	*yoo, oo*
		Past	*taroo, daroo*
	Imperative	—	*ro, e*
non-final clauses	Hypothetical	Provisional	*reba, eba*
		Conditional	*tara, dara*
	Participial	Infinitive	*Ø, i*
		Gerund	*te, de*
		Alternative	*tari, dari*

Although the status of the last three categories in the participial subgroup, i.e. infinitive, gerund, and alternative, is not immediately clear, his classification basically draws on three grammatical criteria, i.e. clause types, mood, and tense, which are universally applicable. Bloch's achievements can thus be recognized as an excellent stepping-stone to further inquiries into the issues of syntax-inflectional morphology correspondences in Japanese. In the rest of this section, I discuss how Bloch's classification can be assimilated into the hierarchical configuration of syntactic structure postulated in generative grammar.

2.3 Generative grammar

One of the important contributions of generative grammar as regards the syntax-morphology relationship is its demonstration that the boundary between syntax

and morphology is not an absolute one. Classical examples of "affix hopping" and "*do* support" in English, exemplified in (1), illustrate this point (Chomsky 1957).

(1) a. John **walked** 5 miles

   b. John **did** not walk 5 miles

   c. Q: **Did** John walk 5 miles?
       A: Yes, he **did** ~~walk 5 miles~~.

These sentences show that the past tense morpheme is syntactically independent in the underlying structure, and that its surface realization is determined by the syntactic environment it appears in. That is, when the tense morpheme T is adjacent to the main verb V, the former gets attached to the latter by affix hopping, as in (1a); when the adjacency is disrupted by intervention of another element as in (1b), movement as in the question in (1c), or deletion as in the answer in (1), the semantically empty auxiliary *do* is inserted to support the morphologically bound tense morpheme. The derivations of the four sentences in (1) are thus schematically represented as in (2a) for (1a), (2b) for (1b) and (1cQ), and (2c) for (1cA).

(2) a. T   V …
         └────────↑

   b. *do*-T   α   V
         ↑ └────X────↑

   c. *do*-T   V̶
         ↑ └──X──↑
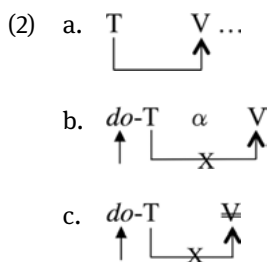
Unless T is treated as an element syntactically independent of V, the contrast between (2a) and (2b, c) and the complementarity between affix hopping and *do* support will not be given a straightforward explanation. More recent theoretical apparatuses such as functional categories (Chomsky 1986, etc.), head movement (Baker 1988, etc.), and Cartography (Cinque (ed.) 2002, etc.) inherit the basic idea that there is no absolute boundary between word-internal and word-external operations, at least as far as inflection is concerned.

The English examples given above thus show that the tense inflection in English is a syntactically independent element that has a node for itself in syntactic structure. Much the same observations can be made with Japanese verbal inflectional endings, where the past tense morpheme is separable from the main verb in certain environments, as shown in (3).

(3)   a.   *tabe-ta*
           eat-PST

    b.   *tabe*   **mo**   *si-ta*
           eat   also   do-PST

(3a) shows that the past tense ending *-ta* is normally attached to the verb stem *tabe-*. In the construction of (3b), on the other hand, the past ending is separated from the main verb by the focus particle *mo* 'also' and thus appears on the support verb *suru* 'do', in parallel with the configuration of the English sentence in (1b). The other inflectional morphemes can also be isolated from the verb in the same environment (*tabe mo su-ru/si-yoo/si-ro/su-reba/si-te*). This observation can be taken as evidence that the Japanese inflectional endings are structurally independent, as they are in English. A caveat should be added here concerning the morphological status of *tabe* in (3). As shown in Table 2, tense and other inflectional endings attach to a verb's stem, whether it is consonant-ending or vowel-ending. The verb *tabe* in (3a) is thus given as a stem. *Tabe* in (3b), however, is not a stem but the *ren'yō* or infinitive form, despite its superficial resemblance. The distinction between a stem and an infinitive manifests itself clearly with consonant-ending verbs like *yom-* 'read'. Compare thus (4a) with (4b).

(4)   a.   *yon-da*   < *yom-ta*
           read [stem]-PST

    b.   *yomi*   *mo*   *si-ta.*
           read [INF]   also   do-PST

In (4a), the past ending *ta*, voiced to *da* under the influence of the final voiced consonant of the preceding morpheme, is attached to the verb stem *yom-*, assimilated to *yon-* because of the dental consonant [t] of *ta*. In (4b), by contrast, the verb 'read' is not given as *yom-* but as *yomi*, which is the infinitive (or nominalized) form of *yom-*. Some morphological adjustments would be called for to make the argument on the Japanese sentences in (3)–(4) strictly parallel to that on the English sentences in (1).

Let us now examine how each of the inflectional endings can be characterized syntactically. There are six inflectional endings that appear in final clauses (i.e. at the clause-ending position), and one of them, the past presumptive form, is morphologically compositional with the past tense morpheme *ta* plus the modal morpheme *roo*. If we assume that T(ense) and M(odal) constitute independent functional categories that project their own phrases, *tabe-taroo* 'will have eaten' will be structurally represented as in (5):

(5)

```
                    MP
              _____|_____
                          M´
                    _____|_____
              TP                M
         _____|_____            |
                  T´           roo
             _____|_____
           V              T
           |              |
         tabe-            ta
```
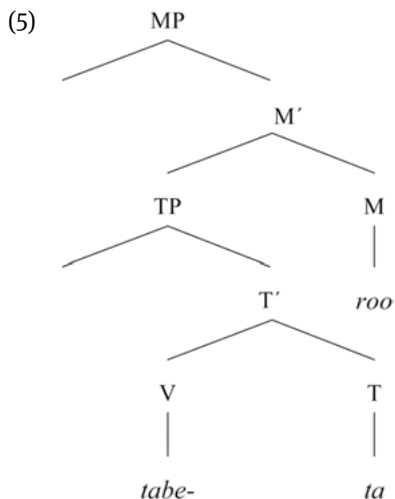
The configurational structure in (5) shows the most transparent case where inflectional endings have a one-to-one correspondence to syntactic categories. The other categories, however, are all morphologically defective in some way or other. The indicative forms contain the overt tense morphemes (*ru* and *ta*) but lack modal marking, and the non-past presumptive form has the visible modal morphology (*(y)oo*) but no tense morphology. Despite the morphological defectiveness, there are syntactic reasons to assume that the morphologically missing elements are in the non-past (or "present") form *-(r)u*. There is not much difficulty finding evidence for this analysis of the indicative and presumptive forms, but more interesting is the imperative form. The imperative ending is *e* with consonant ending verbs (*kak-e*, 'write'), and *ro* with vowel ending verbs (*tabe-ro* 'eat'), and the tense morphology is invisible in both cases. However, the non-past tense morpheme shows up in the case of negative imperatives:

(6) a. *yom-**u***   *na*
         read-NONPST   NEG.IMP
         'Don't read!'

    b. *tabe-**ru***   *na*
         eat-NONPST   NEG.IMP
         'Don't eat!'

In (6), the non-past tense marker *(r)u* appears before the negative imperative marker *na*, which is not an inflectional ending but a sentence-final particle. This clearly shows that Japanese imperatives are tensed. This is further supported by the fact of the following nominative case marking:

(7) *Kimi   ga   ik-e*
     you   NOM   go-IMP
     'You go!'

As Takezawa (1987) argues, nominative case is assigned only in tensed clauses. The overt appearance of a nominative phrase in (7) shows that imperative forms contain tense, although it is morphologically invisible.

Let us move on to the provisional and conditional forms, which Bloch groups together as "hypothetical". They are both tensed, because they can have a nominative subject:

(8) *Taroo   ga   yon-dara / yom-eba,*
     Taro   NOM   read-PROV / read-COND
     'If Taro reads ...'

In view of the fact that these forms are restricted to non-final (i.e. subordinate) clauses, it is most reasonable to assume that they contain some kind of connective function as the head of adverbial clauses that is located at a position higher than T, most likely at C(omplementizer). If so, *reba/eba* and *tara/dara* are amalgamated forms of T and C, although their exact morphological segmentation is yet to be clarified.[1]

We are now left with the last three inflectional categories, infinitive, gerund, and alternative, which Bloch categorizes as "participial". In these cases, the syntactic and semantic function of the alternative form can be identified relatively easily. They clearly have the semantic function of conjoining plural events, and so some kind of conjunction, probably at the level of TP, is involved. If so, *tari* can be analyzed as the morphologically amalgamated form of T and Conj(unction). In contrast, infinitive and gerund forms, which have multiple syntactic functions, are difficult to analyze syntactically in a unified manner. This is especially true in the case of verbs, because these two inflectional verb forms most often get combined with the word after them to form morphologically complex units, and thus their syntactic structures become blurred. On the other hand, infinitive and gerund forms of adjectives and the copula tend to keep their structural independence from the

---

1 Also see Nishiyama (2012) for a syntactic approach to inflection in Japanese. It is interesting to point out here that there are two types of conditional clauses in English; one with the connective *if*, and the other with Subj Aux inversion, or T to C movement in the head movement analysis.

(i) a. [CP If [TP anything should happen to him]], call me at once
     b. [CP Should [TP anything ~~should~~ happen to him]], call me at once.

The fact that in both languages T to C occurs only in hypothetical (conditional/provisional) clauses, not all adverbial clauses (e.g., temporal clauses, reason clauses, purpose clauses, etc.), suggests that something common between English and Japanese is involved in C in hypothetical clauses.

following word, which makes it easier to analyze their structural relations than verbs of the same form. So, in the next section, we first examine inflectional properties of adjectives and the copula, and then in section 5 we analyze the syntactic status of the infinitive and gerund forms by focusing on secondary predicates, in which those two inflectional forms are used.

The suggestions made in this subsection will be concisely summarized in Table 4 with selected examples.

Table 4: Syntactic decomposition of inflectional endings

	STEM	TENSE (ASPECT)	MODAL	COMP (CONJ)
Non-past (or present) indicative	*tabe* *yom*	*ru* *u*		
Past indicative	*tabe* *yon*	*ta* *da*		
Non-past presumptive	*tabe* *yom*		*yoo* *oo*	
Past presumptive	*tabe* *yon*	*ta* *da*	*roo* *roo*	
Conditional	*tabe* *yon*	*ta* *da*		*ra* *ra*
Provisional	*tabe* *yom*		*re* *e*	*ba* *ba*
Imperative	*tabe* *yom*		*ro* *e*	
Alternative	*tabe* *yon*	*ta* *da*		*ri* *ri*
Infinitive	*tabe* *yom*	*i*		
Gerund	*tabe* *yon*	*te* *de*		

Table 4 is intended to show that the generative approach can provide a systematic analysis of Japanese verbal inflections that is not available with traditional Japanese grammar or with Bloch. Japanese inflectional endings are not unanalyzable wholes as assumed by traditional grammarians and Bloch, but are complex amalgams of elements that belong to distinct categories of T, M, and C. These elements, which are configurationally organized in different syntactic positions, are morphologically or phonologically fused with the verb stem, giving a functionally complex ending the superficial appearance of being a single unanalyzable block.

3 Adjectivals

Adjectivals are an especially interesting category in Japanese because there are two types of adjectival words (termed "A" and "AN" in the other chapters of this volume) that have apparently different inflectional properties. In this section, we consider how the two types and their accompanying inflectional endings should be treated syntactically. This section will thus provide discussion supplementary to the discussion on their lexical properties in Chapter 2 (Kishimoto and Uehara, this volume).

Let us begin by observing the basic facts about the two types of adjectives, examples of which are given in (9) and (10).

(9) *Miti   ga   hiro-i / hiro-katta*
     road   NOM   wide-PRS/wide-PST
     'The road is/was wide.'

(10) *Kyoositu   ga   sizuka   da/datta*
      classroom   NOM   quiet   be.PRS/be.PST
      'The classroom is/was quiet.'

hiro- 'wide' in (9) is what is generally referred to as (the stem of) an adjective. This type is traditionally considered to have its own inflectional endings (non-past form *i* and past form *katta*), in much the same way that verbs inflect for tense, as in (11). For this reason, adjectives are included under the rubric of *yōgen* (inflecting words) in the traditional categorization.

(11) *Kodomo   ga   tabe-ru / tabe-ta*
      child   NOM   eat-PRS/eat-PST
      'The child eats/ate.'

On the other hand, *sizuka* 'quiet' in (10) above is followed by *da* and *datta*, which correspond to the non-past and past forms of the copula in the nominal predicate sentence in (12).

(12) *John   ga   sensei   da/datta*
      John   NOM   teacher   be.PRS/be.PST
      'John is/was a teacher.'

This second type is traditionally called "adjectival verbs" (*keiyō-dōshi*), or sometimes called an "adjectival nouns" or "nominal adjectives". In the following, I refer to the first type as Class I, and the second type as Class II to avoid an unnecessary bias. Table 5 shows the entire paradigm of the inflectional endings of the two types of adjectivals.

Table 5: Adjectival inflection

STEM	Class I *taka* 'high'	Class II (copula) *sizuka* 'quiet'
Non-past (present) indicative	*i*	*da*
Past indicative	*katta*	*datta*
Non-past presumptive	*karoo*	*daroo*
Past presumptive	*kattaroo*	*dattaroo*
Imperative	—	—
Provisional	*kereba*	*nara (naraba)*
Conditional	*kattara*	*dattara*
Infinitive	*ku*	*ni* [2]
Gerund	*kute*	*de*
Alternative	*kattari*	*dattari*

The two types indeed have phonetically different endings, but the question is whether or not they should be distinguished syntactically.

The relationship between the two types has been investigated in the generative context as well as in the traditional context. Among traditional grammarians, Hashimoto (1948), who laid the foundation of present school grammar, includes Class I adjectivals in *yōgen* (inflecting words) with their own inflectional endings, whereas Tokieda (1941) claims that they belong in the category of *taigen* (non-inflecting words) with the inflecting copula (*jodōshi* 'auxiliary verb') following them. As discussed in Chapter 2 (Kishimoto and Uehara, this volume), Kageyama (1982) and Miyagawa (1987) attempt to capture their partial similarities and partial differences by amending Chomsky's (1970) feature decomposition of syntactic categories. In the framework of Distributed Morphology (Halle and Marantz 1993), on the other hand, Nishiyama (1999) argues that the two types should be treated as a single syntactic category, with their difference being attributed to the morphophonological irregularities of the copula. We will review Nishiyama's analysis in the rest of this section.

According to Nishiyama, there is no syntactic or semantic difference between the two types of adjectives; it is not that Class I adjectives take their own inflection, but rather they accompany the inflecting copula just like Class II adjectives. Let us look at Class I and Class II more closely. Close inspection of Table 5 suggests a clear parallelism between the two classes; the six inflectional categories (past indicative, non-past and past presumptives, conditional, alternative) exhibit the same morphological compositionality, with the difference only between *ka* and *da* (*katta/datta, karoo/daroo, kattaroo/dattaroo, kattara/dattara, kattari/dattari*). In addition, taking into consideration the endings of verbs in the same six categories (*ta/da, yoo/oo, taro/daroo, tara/dara, tari/dari*), it is most natural to conclude that *ka* and *da* are

---

2 Bloch assumes that the copula lacks the infinitive form and does not deal with *ni*. See Section 4 for discussion on the infinitive status of *ni*.

variants of the same copula and the rest of the morphemes are inflectional endings. Table 5 shows some irregular forms involving the Class I non-past form *i*, but such an irregularity is common in inflectional morphology across languages, especially in the case of functional verbs like the copula (see Nishiyama 1999 for more discussion).

As against our contention that Class I and Class II adjectives should be analyzed as belonging to the same category, both taking the copula, some pieces of syntactic evidence have been adduced that the two types of adjective behave differently. One piece of the oft-cited evidence is the adjectives' compatibility with modal predicates like *mitai-da* 'seem', *rasi-i* 'be likely' and *daroo* 'probable/seem'. Miyagawa (1987), for instance, claims, on the basis of the following data, that the Class I adjective *taka* 'high' and the Class II adjective *kirei* 'pretty' behave differently with respect to the co-occurrence with *mitai-da* and hence do not form a natural class; he argues, rather, that Class I forms a natural class with verbs and Class II, with nouns:

(13) a. **taka* *mitai-da* (Class I) b. *kirei* *mitai-da* (Class II)
          high   seem-COP   pretty seem-COP

     c. **tabe* *mitai-da* (verbs) d. *otoko* *mitai-da* (predicate nominal)
          ea   seem-COP   man   seem-COP

However, Nishiyama points out that (13) is a pseudo-paradigm; a different picture emerges if we take into account the following data including tense inflection:

(14) a. *{tabe-ru / tabe-ta}* *mitai-da* (verbs)
          {eat-PRST / eat-PST} seem-COP

     b. *{taka-i / taka-katta}* *mitai-da* (Class I)
          {high-PRS / high-PST} seem-COP

     c. *{kirei-datta / *kirei-da}* *mitai-da* (Class II)
          {pretty-be.PST / *pretty-be.PRS} seem-COP

     d. *{otoko-datta / *otoko-da}* *mitai-da* (predicate nominal)
          {man-be.PST / *man-be.PRS} seem-COP

The paradigm in (14) shows that *mitai-da* selects complement clauses with a tensed predicate (both non-past and past), with the exception of the non-past copula *da* in (14c, d). In other words, the real question is not about the category matching of *mitai-da* with Class II adjectives and nouns or with Class I adjectives and verbs, as Miyagawa argues, but the non-occurrence of *da* in the complement clause of *mitai-da*, which takes a TP, or tensed clause, as its complement. In addition, the complex form of the copula, *de ar(u)* (predicative copula + dummy copula), which Nishiyama

claims to be amalgamated to the monomorphemic *da*, following Nakayama (1988) and Urushibara (1993), can appear with *mitai-da*:

(15) a. *kirei-de*   *ar-u*   *mitai-da*
     pretty-PRED.COP   DUM.COP.PRS   seem-COP

    b. *otoko-de*   *ar-u*   *mitai-da*
     man-PRED.COP   DUM.COP.PRS   seem-COP

This means that what is involved here is not a general issue concerning syntactic categories, but a particular problem pertaining only to the deletability of the simple, non-past form of the copula *da*.

Given the peculiarity of *da*, it is expected that this particular form causes exceptional behavior in other syntactic environments. In fact, it is observed that only *da*, but not the past form *datta* nor the complex nonpast form *de ar(u)*, can be deleted before sentence particles like *yo* (assertive particle), *ne* (emphatic particle), and in some other contexts. This also indicates that the phenomena in question should be isolated from the discussion of the issue of categorial distinction.

In this section, we have argued, following Nishiyama (1999), that it is not necessary to posit two distinct classes of adjectives based on syntactic criteria, and that Class I and Class II adjectives differ only in the morphophonological properties of the copula that follows them.

4 Infinitives, gerunds, and secondary predicates

The various kinds of inflections discussed so far have the function of closing off finite sentences or non-finite subordinate clauses. In this section, we tackle another manifestation of inflectional categories in two types of secondary predication, namely resultative predicates and depictive predicates. By looking into the nature of the two inflectional forms which Bloch calls infinitive and gerund, we aim to shed new light on the relationship between syntax and inflectional morphology.

Resultative predicates, as exemplified in (16), describe the final state that the theme argument of the main predicate reaches as a result of the event denoted by the main verb; depictive predicates, as in (17), on the other hand, delineate the state of the subject or object at the time when the event expressed by the main predicate happens (Rothstein 1983).

(16) Resultative predicates
    a. *Sakana*   *ga*   ***ooki-ku / kyodai-ni***   *sodat-ta*
      fish   NOM   big-KU / huge-NI   grow-PST
      'The fish grew *big/huge*.'

b. *Taroo   ga   kabe   o   **aka-ku / makka-ni**   nut-ta*
    Taro   NOM   wall   ACC   red-KU / crimson-NI   paint-PST
    'Taro painted the wall *red/crimson*.'

(17) Depictive predicates
      a. *Taroo   ga   **tukare-te / hadaka-de**   kaet-te   ki-ta*
          Taro   NOM   get.tired-TE / naked-DE   return-GER   come-PST
          'Taro came home *tired / nude*.'

      b. *Taroo   ga   niku   o   **nama-de**   tabe-ta*
          Taro   NOM   meat   ACC   raw-DE   eat-PST
          'Taro ate the meat *raw*.'

As shown by the boldfaced words in (16) and (17), the resultative and depictive predicates in Japanese are coded by different inflectional endings.

There is abundant literature dealing with the syntactic and semantic properties of resultative and depictive constructions, such as Takezawa (1991), Kageyama (1996) and Washio (1997) for resultatives, and Koizumi (1994) and Takezawa (2002) for depictives. The morphological status of these two kinds of predicate, however, has eluded due attention. Below we explore their morphological status from the viewpoint of inflection.

4.1 Resultative predicates

As shown in (16) above, resultative predicates end in *ku* and *ni*, which are the infinitive forms of the copula for Class I and Class II adjectives, respectively (cf. Table 5). As argued in Section 3, the two forms of inflectional endings, including *ku* and *ni*, are regarded as allomorphs of the same copula. Since they have the same syntactic status, it is not at all surprising that they can both appear as resultative predicates (see also Okutsu 1978). If they have different categorial specifications, as argued by Kageyama (1982) and Miyagawa (1987), then an explanation is required for the fact that they behave identically in this particular syntactic context.

The infinitive forms of the copula *ku/ni* also show up in syntactic environments other than resultative predicates. They appear in complements to the verb of thinking *omo(u)* 'think' in (18a), the causative verb *su(ru)* 'make' in (18b), and the change-of-state verb *nar(u)* 'become' in (18c).

(18) a. *Taroo   wa   Hanako   o   utukusi-**ku** / kirei-**ni**   omot-ta*
          Taro   TOP   Hanako   ACC   beautiful-KU / pretty-NI   think-PST
          'Taro thought Hanako to be beautiful/pretty.'

b. *Hanako   wa   kabe   o   aka-**ku** / makka-**ni**   si-ta*
   Hanako   TOP   wall   ACC   red-KU / crimson-NI   make-PST
   'Hanako made the wall red / crimson.'

c. *Kabe   ga   aka-**ku** / makka-**ni**   nat-ta*
   wall   NOM   red-KU / crimson-NI   become-PST
   'The wall became red / crimson.'

As discussed by Takezawa (1987) and Takezawa and Whitman (1998), complements to these verbs correspond to the Exceptional Case Marking and raising types of complements in English, and can thus be characterized as defective infinitive clauses. A piece of evidence for this comes from the fact that nominative case cannot occur in these complements. In (18a, b), the thematic subjects of the embedded clauses, *Hanako* and *kabe*, cannot take nominative *ga* but must have accusative *o*. Taking these observations into account, it is plausible to identify *ku* and *ni* as tenseless copulas without any nominative assigning/checking ability. Lack of nominative checking allows them to occur only in infinitive complements, as well as resultative predicates. The term "infinitive", as Bloch defines it, is deemed appropriate for the inflectional category of *ku* and *ni*.

How, then, are infinitive clauses with *ku* and *ni* structurally represented? Nishiyama (1999) claims that the copula projects itself to form a predicative phrase, which he calls a PredP, following the proposal by Bowers (1993). Nishiyama further argues that *ku* and *ni* infinitives are licensed if the PredP is directly dominated by the VP, as illustrated below.

(19)
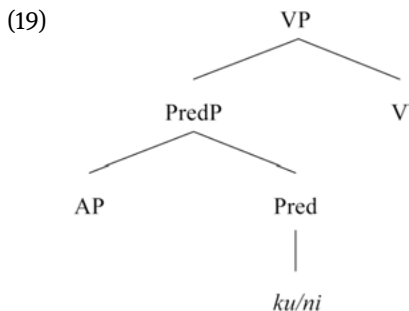
There is, however, another grammatical usage of the infinitive form of the copula. The *ku/ni* infinitive forms also function as manner adverbials, as shown in (20).

(20) a. *Taroo   wa   haya-ku / asibaya-ni   arui-ta*
        Taro   TOP   fast-KU / fast-NI   walk-PST
        'Taro walked fast.'

   b. *Hanako   wa   uma-ku / zyoozu-ni   utat-ta*
      Hanako   TOP   good-KU / good-NI   sing-PST
      'Hanako sang well.'

Namai (2002) criticizes this line of analysis, stating that the predicate analysis of *ku/ni* does not account for this adverbial usage because manner adverbs are not considered predicates. Nishiyama (2005) replies to this criticism by saying that manner adverbs can be regarded as a kind of predicate in the sense that they are predicated over events. If so, the structure in (19) can give a unified account of the occurrence of the *ku/ni* infinitive forms as manner adverbs, because manner adverbs are widely assumed to take VP-adjoined positions. Since adverbs clearly do not contain tense specification, this structure could accommodate the adverbial usage of the infinitive forms as well.[3]

4.2 Verbal depictive predicates

Japanese has two types of depictive predicates, as shown in (21), which are repeated below:

(21) a. *Taroo   ga   tukare-**te** / hadaka-**de**   kaetteki-ta*
     Taro   NOM   get.tired-TE / naked-DE   return-PST
     'Taro came back *tired* / *nude*.'

   b. *Taroo   ga   niku   o   nama-**de** tabe-ta*
     Taro   NOM   meat   ACC   raw-DE   eat-PST
     'Taro ate the meat *raw*.'

The first type involves the inflectional verb form which Bloch calls "gerund". *Tukare-te* 'tired' in (21a) is an example of this type. I call this type the *V-te* form below for the reason to be stated shortly. The second type involves expressions like *hadaka-de* 'naked' (21a) and *nama-de* 'raw' (21b), which I call the *X-de* form.

   Let us consider the *V-te* type, first. More examples of *V-te* depictive predicates are given in (22):

(22) a. *Taroo   ga   sofa   de   yopparat-te   ne-ta*
     Taro   NOM   sofa   LOC   get.drunk-TE   fall.in.sleep-PST
     'Taro fell in sleep on the sofa *drunk*.'

   b. *Hanako   ga   paatii   ni   kikazat-te   syussekisi-ta*
     Hanako   NOM   party   LOC   dress.up-TE   attend-PST
     'Hanako went to the party *dressed up*.'

---

3 Nishiyama's analysis of manner adverbs implies that the alternation between adjectives and adverbs is not a derivational conversion, but is to be treated as inflectional in Japanese (cf. Baker 2003; Takezawa 1991). This poses an interesting question for the investigation of the relationship between derivational and inflectional morphology.

  c. *Taroo ga uti kara okot-te   tobidasi-ta*
    Taroo NOM house ABL get.angry-TE run.out-PST
    'Taro ran out of the house *angry*.'

Notice here that the expressions used as depictive predicates, corresponding to the adjectives in English, are all unaccusative verbs, expressing change of state (*yopparaw(u)* 'get drunk', *okor(u)* 'get angry', *kikazar(u)* 'dress up'). In addition, it is commonly known in Japanese linguistics (Kindaichi 1950) that they express resultant states when they appear with the auxiliary verb *i(ru)* 'be' in primary predication, as in (23).

(23) *Taroo ga yopparat-te / kikazat-te / okot-te   i-ru*
    Taro NOM get.drunk-TE / dress.up-TE / get.angry-TE be-PRS
    'Taro is tired / drunk / angry.'

In contrast, when unergative and transitive verbs expressing action, such as *hasir(u)* 'run' and *hur(u)* 'wave', are used in the *V-te* form in the same syntactic environment, they denote progressive actions occurring simultaneously with the actions expressed by the main verbs. They are roughly synonymous with the aspectual adverbial expression *nagara* 'while', as shown in (24).

(24) a. *Taroo ga uti made hasit-te / hasiri-nagara kaet-ta*
    Taro NOM home to run-TE / run-while return-PST
    'Taro returned home, running.'

  b. *Taroo ga sensyu o hata o hut-te / huri-nagara ooensi-ta*
    Taro NOM players ACC flags ACC wave-TE / wave-while cheer -PST
    'Taro cheered the players, waving a flag.'

These unergative and transitive verbs have progressive interpretation when followed by the auxiliary *i(ru)* in primary predication.

(25) a. *Taroo ga   hasit-te i-ru*
    Taro NOM run-TE be-PRS
    'Taro is running'

  b. *Taroo ga   hata o   hut-te i-ru*
    Taro NOM flags ACC wave-TE be-PRS
    'Taro is waving a flag'

This parallelism demonstrates that the inflectional ending *te* in *V-te* depictive predicates and *te* in the *V-te i(ru)* sequences are the same elements. It can also be concluded that although the *V-te i(ru)* sequence is often treated as a unit that carries

an aspectual interpretation as a whole, it is just the inflectional ending *te* that is the real locus of the aspectual interpretation, the auxiliary verb *i(ru)* being a dummy verb supporting tense (Uchimaru 2006, but see Chapter 11 [Nakatani, this volume] for a different analysis). Given the discussion so far, it is reasonable to assume that depictive predicates occur in the following syntactic structure, where *te* is the head of the functional category Asp and the AspP functions as an adjunct expressing the state of the subject in the matrix clause through predication (or control):

(26)
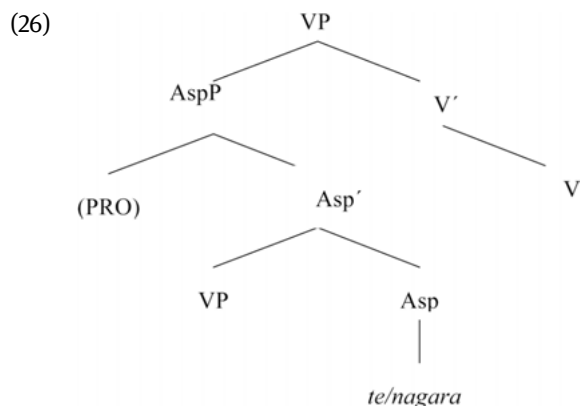
4.3 *X-de* depictive predicates

The second type of depictive has the form *X-de*. So far we have not identified the syntactic status of *X* and *de*. This section examines the syntactic status of *X* first, and then the next section analyzes *de* based on the discussion in this section.

In (21) above, we saw a couple of depictive expressions of the *X-de* type. Some more illustrative examples are added below:

(27) a. *Taroo   ga   reesu   ni   hadaka-de   sankasi-ta*
Taro   NOM   race   DAT   naked-DE   participate-PST
'Taro participated in the race *naked*.'

  b. *Taroo   ga   sakana   o   nama-de   tabe-ta*
Taro   NOM   fish   ACC   raw-DE   eat-PST
'Taro ate the fish *raw*'

  c. *Basu   ga   taaminaru   ni   kara-de / man'in-de   tootyakusi-ta*
bus   NOM   terminal   LOC   empty-DE / full-DE   arrive-PST
'The bus arrived at the terminal *empty / full*.'

Considering the correspondence with the English counterparts, the *X* in the *X-de* type, such as *hadaka* 'naked', *nama* 'raw', *kara* 'empty', and *man'in* 'full', might be

taken to belong to Class II adjectives, and the *de* following X, as the gerund form of the copula *da*. In fact, Okutsu (1997) suggests this possibility. However, there is an interesting contrast between (27) and (28), both of which have the same *X-de* form:

(28) a. **Hanako   ga   paatii   ni   hanayaka-de   sankasi-ta*
    Hanako   NOM   party   LOC   gorgeous-DE   participate-PST
    'Hanako went to the party gorgeous'

   b. **Taroo   ga   mikan   o   sinsen-de   tabe-ta*
    Taro   NOM   orange   ACC   fresh-DE   eat-PST
    'Taro ate the orange *fresh*'

   c. **Basu   ga   taaminaru   ni   sizuka-de   tootyakusi-ta*
    bus   NOM   terminal   LOC   quiet-DE   arrive-PST
    'The bus arrived at the terminal *quiet*'

Although (27) and (28) share the same *X-de* form superficially, there is a clear difference in acceptability between them. Interestingly, this difference can be attributed to the categorial status of *X*, as pointed out by Takezawa (2002) (cf. Nishiyama 1999). The words corresponding to *X* in (27) appear to have adjectival meanings which describe properties of persons/things. Although the English counterparts of *X* in (27) are all adjectives, in Japanese they are categorized as nouns. The categorial difference of *X* between (27) and (28) can be shown by a number of syntactic tests. The first test is concerned with the form of prenominal modification. When a noun modifies a noun, the first noun is marked with *no*, whereas when a Class II adjective modifies a noun, it is marked with *na*:

(29) a. *hadaka-{no / *na} otoko* [naked-{no / *na} man]

   b. *nama-{no / *na} sakana* [raw-{no / *na} fish]

   c. *man'in-{no / *na} basu* [full-{no / *na} bus]

(30) a. *hanayaka-{na / *no} zyosei* [gorgeous-{na / *no} lady]

   b. *sinsen-{na / *no} sakana* [fresh-{na / *no} fish]

   c. *sizuka-{na / *no} basyo* [quiet-{na / *no} place]

The second test concerns the nominalizing suffix *-sa*. Nouns cannot undergo *sa*-nominalization, whereas adjectives can:

(31) **hadaka-sa* [naked-NMLZ], **nama-sa* [raw-NMLZ], **man'in-sa* [full-NMLZ]

(32) *hanayaka-sa* [gorgeous-NMLZ], *sinsen-sa* [fresh-NMLZ], *nigiyaka-sa* [noisy-NMLZ]

These two tests clearly indicate that depictive predicates of the *X-de* type must contain a noun, not a Class II adjective. Note further that Class I adjectives can never appear as depictive predicates, regardless of whether the ending is the gerund *kute* form or the infinitive *ku* form:

(33) a. **Hanako   ga   paatii   ni   utukusi-ku / -kute   sankasi-ta*
        Hanako   NOM   party   LOC   beautiful-KU / -KUTE   participate-PST
        'Hanako went to the party beautiful'

     b. **Taroo   ga   sakana   o   tumeta-ku / -kute   tabe-ta*
        Taro   NOM   fish   ACC   cold-KU / -KUTE   eat-PST
        'Taro ate the fish cold'

     c. **Basu   ga   taaminaru   ni   soozoosi-ku / -kute   tootyakusi-ta*
        bus   NOM   terminal   LOC   noisy-KU / -KUTE   arrive-PST
        'The bus arrived at the terminal noisy'

This demonstrates not only that it is the syntactic category, and not the morphological shape, that restricts the occurrence of depictive predicates, but also that Class I and Class II adjectives form one natural class.

So far we have argued that the *X* in *X-de* type depictives must be a noun and cannot be an adjective, whether Class I or Class II. Note in passing a difference in categorization between Japanese and English. We saw that in English, *raw, naked, empty*, etc. belong to the category of adjective, whereas the corresponding Japanese words, *nama, hadaka, kara*, etc. are all nouns. What creates this difference between the two languages? What is the criterion that divides nouns and adjectives in each of the two languages? It has been pointed out that "gradability" is an important factor that separates the two categories in Japanese (Uehara 1998). As illustrated in (34), adjectives, whether Class I or II, can be modified with degree adverbs such as *sukosi* 'slightly', *totemo* 'very', *motto* 'more', but the class of words that appear as depictives with *de* cannot be modified in this way.

(34) a. *{sukosi / totemo / motto} {utukusi-i / tumeta-i / hanayaka-da / sinsen-da}*
        {slightly / very / more}   {beautiful-PRS / cold-PRS / gorgeous-PRS / fresh-PRS}

     b.**{sukosi / totemo / motto} {nama-de / hadaka-de / kara-de / ippai-de}*
        {slightly / very / more}   {raw-DE / naked-DE / empty-DE / full-DE}

In contrast, gradability does not constitute the absolute line dividing these two categories in English, and hence English has many non-gradable adjectives.

A group of words that shows interesting properties with respect to the noun/ adjective contrast is mimetic words. Mimetics basically have two usages, adverbial usage and predicative usage. The former optionally accompanies the particle *to* and

describes manners of dynamic activity, as in *tobotobo(-to) aruku* 'walk wearily', *surasura(-to) hanasu* 'speak fluently', *(ame-ga) zaazaa(-to) huru* 'to rain hard'. The latter generally take the copula and describe static states, as in *hetoheto da* [exhausted COP], *betobeto da* [sticky COP], and *kotikoti da* [hard COP].[4] What interests us here is the predicative usage. Mimetic words look like Class II adjectives when used in primary predication like in (35), as they describe properties/states of things:

(35) a. *Sitai   ga   **barabara**   da*
       body   NOM   dismembered   COP.PRS
       'The body was dismembered.'

    b. *Basu   ga   **garagara**   da*
       bus   NOM   nearly.empty   COP.PRS
       'The bus is nearly empty.'

    c. *Sensyu   ga   **hetoheto**   da*
       player   NOM   exhausted   COP.PRS
       'The player is exhausted.'

However, they can be used as depictive predicates with *de*:

(36) a. *Sitai   ga   **barabara-de**   mitukat-ta*
       body   NOM   dismembered-DE   be found-PST
       'The body was found dismembered.'

    b. *Basu   ga   taaminaru   ni   **garagara-de**   tootyakusi-ta*
       bus   NOM   terminal   LOC   nearly.empty-DE   arrive-PST
       'The bus arrived at the terminal nearly empty.'

    c. *Sensyu   ga   guraundo   kara   **hetoheto-de**   hikiageteki-ta*
       player   NOM   ground   ABL   exhausted-DE   withdraw-PST
       'The player withdrew from the ground *exhausted*.'

This suggests that mimetic words are categorially nouns. In fact, they exhibit nominal properties in the tests for distinguishing nouns and adjectives:

---

4 See Suzuki (2012) for discussion on the predicative use of mimetic words with a focus on the three different predicative forms, X *da* 'be X', X *suru* 'do X', and X *site iru* 'be doing X'. For example, the mimetic word *turuturu* 'smooth/slippery' takes all three forms, *turusuru da, turuturu suru*, and *turuturu site iru*.

(37) Selection of *no*

    a.   *barabara-{no / *na}*   *sitai*
          dismembered-*{no / *na}*   body

    b.   *garagara-{no / *na}*   *basu*
          nearly.empty-*{no / *na}*   bus

    c.   *hetoheto-{no / *na}*   *sensyu*
          exhausted-*{no / *na}*   player

(38) Inapplicability of *sa*-nominalization

     *barabara-sa   **garagara-sa*   **hetoheto-sa*
     dismembered-NMLZ   nearly.empty-NMLZ   exhausted-NMLZ

(39) Degree modification

     *{sukosi / totemo}   *{garagara / barabara / hetoheto}*   *da*
     {slightly / very}   {nearly.empty / dismembered / exhausted}   COP.PRS

Therefore, mimetics belong to the category of nouns, even though they appear to be adjectival in meaning, and that is why they can be used as depictive predicates with *de*.

     In this section, we have examined the *X-de* type depictive predicates and demonstrated that X must a noun, although the corresponding English expressions are adjectival. Furthermore, we have shown that mimetic words are semantically property-describing, but they are categorically nouns and hence can be used in depictive expressions.

4.4 The status of *de* in the X-*de* type depictives

In this section, we focus on the syntactic status of *de*. As mentioned in the previous section, one of the possibilities of the status of *de* is the *te*-form ("gerund" in Bloch's terminology) of the copula (Okutsu 1997; Nishiyama 1999). If *X* is a Class II adjective, this is the only possibility, because adjectives must be followed by the copula in Japanese. However, given that the *X* in X-*de* type depictives is actually a noun, there is another possibility, namely that *X* is a postposition. In what follows, we argue that *de* should be analyzed not as the *te*-form of the copula but as a P, by comparing the X-*de* and V-*te* forms.

     Before comparing the *V-te* type and the *X-de* type, let us consider the following paradigm of data involving the perception verb *mie(ru)* 'look (like)':

(40) a. *Taroo ni wa sono miti ga massugu-{ni / *de} mie-ru*
Taro DAT TOP that road NOM straight-{NI / *DE} look-PRS
'The road looks straight to Taro'

b. *Taroo ni wa sono miti ga hiro-{ku / *kute} mie-ru*
Taro DAT TOP that road NOM wide-{KU / *KUTE} look-PRS
'The road looks wide to Taro'

c. *Taroo ni wa sono miti ga kawa {ni / *de} mie-ru*
Taro DAT TOP that road NOM river {NI / DE} look-PRS
'The road looks like a river to Taro'

d. *Taroo ni wa miti ga {magat-te / *magari} mie-ru*
Taro DAT TOP road NOM {bend-TE / bend.INF} look-PRS
'The road looks winding to Taro'

Here, we find an interesting difference in the form of embedded predicates between (40a–c) and (40d). *Mie(ru)* takes the *ni/ku* infinitive form when the predicate is copulative, as in (40a–c), while it takes the *te*-form (or gerund form) when it is verbal, as in (40d). Recall that *ku* and *ni* are allomorphs of the same copula and syntactically non-distinct.

What causes this difference between copulative and verbal predicates? We assumed in Section 4.2 that the *te* is the head of the AspP that bears the function of determining the resultative or progressive interpretation depending on the aspectual properties of the main verb to which it is attached. If we assume that the perception verb *mie(ru)* requires stative predicates, the aspectual morpheme *te* must be present to turn the complement verb to such a predicate. On the other hand, a copulative predicate, whether it involves an adjective or a noun, is inherently stative, and therefore the *te* form of the copula (*de* for Class I and *kute* for Class II) cannot appear in the complement of the perception verb.

If this account of the paradigm (40) is correct, we cannot treat *te* in *V-te* type depictives and *de* in *X-de* type depictives in the same way. This leads us to conclude that *de* in the *X-de* type is not the aspectual *te*-form, but something else. The remaining possibility is that it is a postposition.

There is another phenomenon in which the *V-te* and *X-de* differ.

(41) a. *Taroo_i wa paatii ni Hanako_j o yopparat-te_{i/*j} ture-te it-ta.*
Taro TOP party LOC Hanako ACC get.drunk-TE take-GER go-PST
'Taro took Hanako to the party drunk'

b. *Taroo_i wa paatii ni Hanako_j o hutukayoi-de_{i/j} ture-te it-ta*
Taro TOP party LOC Hanako ACC hang.over-DE take-GER go-PST
'Taro took Hanako to the party with a hangover'

c. *Taroo$_i$ wa paatii ni Hanako$_j$ o yopparatta mama-de$_{i/j}$ tureteit-ta*
   Taro TOP party LOC Hanako ACC drunk   while-DE take-PST
   (lit) 'Taro took Hanako to the party in a drunk state'

(41a) contains a *V-te* type depictive *yopparat-te* 'get drunk-TE', (41b), a *X-de* type *hutukayoi-de* 'hang over-DE', and (41c), the abstract noun *mama + de*. Here, the *V-te* forms in (41a) can be the predicate of the subject, but not of the object, whereas *X-de* in (41b,c) can be the predicate of both the subject and object. These two types also differ in this respect, and thus *de* in *X-de* cannot be analyzed as a *te*-form of the copula. The possibility left for *de* is that it is a P.

In this subsection, we have compared *X-de* type depictives with *V-te* depictive and demonstrated that they behave differently, thereby concluding that *de* in the *X-de* cannot be identified with the *te* form of the copula, but should be analyzed as a postposition.

4.5 Summary

Section 4 has examined categorial and morphological characterizations of resultative and depictive secondary predicates in Japanese. The forms of each secondary predicate are tabulated in Table 6:

Table 6: Forms of resultative and depictive predicates

	adjective (+copula)	noun (+copula)	verb
resultative	Class I: A-*ku* (infinitive) Class II: A-*ni* (infinitive)	N-*ni* (infinitive)	*
depictive	*	N-*de* (P)	V-*te* (gerund)

There are two gaps in this table, verbal resultatives and adjectival depictives. At this point, I have no interesting account of why they do not exist in Japanese, but deeper investigation into the question might give some clue to understanding the relation between syntactic categories and morphological forms of secondary predicates in Japanese, and predicate forms in general.[5]

5 Conclusion and future research perspectives

The study of inflection in Japanese has a long history, but most of the literature in traditional Japanese grammar is focused on its morphophonological and semantic

---

5 It might be the case that V V compounds like *hasiri tukareru* [run get.tired] fill the gap of verbal resultatives. See Hasegawa (2000) for related discussion.

properties. In this chapter, drawing on the inflectional paradigm offered by Bernard Bloch, I have attempted to characterize the Japanese inflectional system in terms of syntactic categories and syntactic structures assumed in generative grammar. The major claims of the chapter will be summarized as follows:

- The ten inflectional suffixes of verbs Bloch identified by the structuralist method can be given systematic characterizations syntactically in terms of functional categories such as C, M, T, and some others (Table 4).
- There is no categorial difference between Class I adjectives (*utukusi-i* 'beautiful') and Class II adjectives (*kirei-da* 'pretty'), both of which can be analyzed as taking the inflecting copula after them.
- The two types of non-finite secondary predicates, resultative and depictive, have different categorical and inflectional properties, unlike their English counterparts; the former require the infinitive (*ku/ni*) forms of the copula, whereas the latter take the Noun-Postposition form or the gerund form of verbs.

While we argued, following Nishiyama (1999), that the inflectional endings of both Class I and Class II adjectives are variants of copulas, there are other potential accounts whose feasibility needs to be examined closely. One is the strong lexicalist approach like Sells (1995) in which all the inflectional endings of the two classes are uniformly treated as lexical suffixes (see Koopman (2005) for arguments against Sells (1995)). Another is Yamakido (2005), where, based on dialectal, historical, and typological data, a novel proposal is developed that the adjectival inflections in both *-i* and *-na* paradigms are identified as case-markers. While all these analyses aspire to a unitary treatment of Class I (i.e. true adjectives) and Class II (i.e. adjectival nouns), evidence from Chapter 14 (Kageyama, this volume) suggests that the endings in the *-i* paradigm and those in the *na* paradigm are morphologically distinct, the former being morphological suffixes and the latter clitics or *fuzoku-go* (non-independent words). The properties of the conjugational endings accompanying the two classes of adjectives are highly complex because of heavy fusion and elision of inflectional morphemes in their historical development, so much so that their status could be best regarded as an unsettled issue.

The discussion in this chapter covers only limited portions of Japanese inflectional phenomena and there are many problems I could not deal with here. One of those problems which is closely associated with the issues discussed above is the *rentaikei* (adnominal forms of the copula, that is, the selection between *na* and *no* in prenominal environments, as in *sinsen-na/*no yasai* [fresh-COP vegetable] and *nama-no/*na yasai* [raw-COP vegetable]. The fact that the adnominal forms exhibit phonetically distinct variants only with the copula *da* in Modern Japanese raises a question of whether they have any syntactic effect in synchronic grammar or are mere morphological variants (Hiraiwa 2001). Another related question is how to treat the copula *da* and its morphological variants, the complex form *de-ar(-u)* and the polite form *des(-u)*. They are used not just as markers for predication, but also

serve for other functions such as style and politeness markers, sometimes exhibiting different distributional properties among them (see Teramura 1984; Okutsu 1978; among others). These questions are related to one another and need to be investigated in detail in order to provide a unified account of the interaction between syntax and inflectional morphology in Japanese.

Acknowledgments

References

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark. 2003. *Lexical categories: Verbs, nouns, and adjectives*. Cambridge: Cambridge University Press.

Bloch, Bernard. 1946a. Studies in colloquial Japanese I: Inflection. *Journal of the American Oriental Society* 66. 97–109. Reprinted in Roy A. Miller (ed.), 1969. *Bernard Bloch on Japanese*. New Haven: Yale University Press.

Bloch, Bernard. 1946b. Studies in colloquial Japanese II: Syntax. *Language* 22. 200–248. Reprinted in Roy A. Miller (ed.). 1969. *Bernard Bloch on Japanese*. New Haven: Yale University Press.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24(4). 591–656.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In Rodney A. Jacobs and Peter A. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Cinque, Guglielmo (ed.). 2002. *Functional structure in DP and IP: The cartography of syntactic structures, volume 1*. New York: Oxford University Press.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale and Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176, Cambridge, MA: MIT press.

Hasegawa, Nobuko. 2000. Resultatives and language variations: Result phrases and VV compounds. In Mineharu Nakayama and Charles J. Quinn, Jr. (eds.), *Japanese/Korean Linguistics* 9, 269–282. Stanford: CSLI Publications.

Hashimoto, Shinkichi. 1948. *Shin bunten bekki kōgo hen*. [New grammar of colloquial Japanese] Tokyo: Fuzambō.

Hiraiwa, Ken. 2001. On Nominative-genitive conversion. In Elena Guerzoni and Ora Matushansky (eds.), *A few from Building E39 (MIT Working Papers in Linguistics 39)*, 65–123. Cambridge, MA: MITWPL.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1996. *Dōsi imiron: Gengo to ninti no setten* [Semantics of verbs: Interface between language and cognition]. Tokyo: Kurosio Publishers.

Kageyama, Taro. this volume. Chapter 14: Lexical integrity and the morphology-syntax interface. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation.* Berlin: De Gruyter Mouton.

Kindaichi, Haruhiko. 1950. Kokugo dōshi no ichibunrui [A classification of Japanese verbs]. *Gengo Kenkyu* 15. 41–63.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation.* Berlin: De Gruyter Mouton.

Kinsui, Satoshi. 1997. Kokubunpō [Traditional Japanese grammar]. In Takashi Masuoka and Yoshio Nitta (eds.), *Iwanami kōza, Gengo no kagaku 5: Bunpō*, 119–157. Tokyo: Iwanami Shoten.

Koizumi, Masatoshi. 1994. Secondary Predication. *Journal of East Asian Linguistics* 3(1). 25–79.

Koopman, Hilda. 2005. Korean (and Japanese) morphology from a syntactic perspective. *Linguistic Inquiry* 36(4). 601–633.

Miyagawa, Shigeru. 1987. Lexical categories in Japanese. *Lingua* 73. 29–51.

Mikami, Akira. 1963. *Nihongo no kōbun* [Constructions in Japanese]. Tokyo: Kurosio Publishers.

Nakatani, Kentaro. this volume. Chapter 11: Complex predicates with *-te* participial verbs. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation.* Berlin: De Gruyter Mouton.

Nakayama, Mineharu. 1988. Empty copulas. In Yukio Otsu (ed.), *Mita working papers in psycholinguistics* 1. 121–128. Tokyo: Keio University,

Namai, Ken-ichi. 2002. The word status of Japanese adjectives. *Linguistic Inquiry* 33(2). 340–349.

Nishiyama, Kunio. 1999. Adjectives and the copulas in Japanese. *Journal of East Asian Linguistics* 8(3). 183–222.

Nishiyama, Kunio. 2005. Morphological boundaries of adjectives: A reply to Namai. *Linguistic Inquiry* 36(1). 134–143.

Nishiyama, Kunio. 2012. Katsuyōkei no keitairon, tōgoron, on'inron, tsūji [Morphology, syntax, phonology, and diachrony of inflectional forms]. In Ken'ichi Mihara and Nitta Yoshio (eds.), *Katsuyōron no zensen* [Frontiers of Inflectional morphology], 153–189. Tokyo: Kurosio Publishers.

Okutsu, Keiichiro. 1978. *"Boku wa unagi da" no bunpō* [The grammar of "Boku wa unagi da"]. Tokyo: Kurosio Publishers.

Okutsu, Keiichiro. 1997. Rentai soku ren'yō? [Are adnominals adverbals?]. *Nihongogaku* 16–10.

Rothstein, Susan. 1983. *The syntactic form of predication*. Cambridge, MA: MIT dissertation.

Sadakane, Kumi and Masatoshi Koizumi. 1995. On the nature of the "dative" particle *ni* in Japanese. *Linguistics* 33(1). 5–33.

Sakuma, Kanae. 1936. *Gendai nihongo no hyōgen to gohō* [Expressions and grammar of Modern Japanese]. Tokyo: Kōseikaku.

Sells, Peter. 1995. Korean and Japanese morphology from a lexical perspective. *Linguistic Inquiry* 26(3). 237–325.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Suzuki, Ayaka. 2012. Nihongo on onomatope jutsugo no keisiki nitsuite [On the forms of mimetic predicates in Japanese]. *Nihongo Bunpō* [Journal of Japanese Grammar] 12(2). 162–178.

Takezawa, Koichi. 1987. *A configurational approach to Case marking in Japanese*. Seattle: University of Washington dissertation.

Takezawa, Koichi. 1991. Secondary predication and goal/locative phrases. In Hasegawa Nobuko (ed.), *Japanese syntax in comparative grammar*, 45–87. Tokyo: Kurosio Publishers.

Takezawa, Koichi. 2002. Nihongo no jōtai kijutsu nijijutsubu to tōgo hanchū [Secondary depictive predicates and syntactic categories in Japanese]. *Proceedings of Kansai Linguistic Society* 22. 257–267.

Takezawa, Koichi and John B. Whitman. 1998. *Kaku to gojun to tōgo kōzō* [Case, word order, and syntactic structure]. Tokyo: Kenkyusha.

Teramura, Hideo. 1984. *Nihongo no shintakusu to imi II* [Syntax and semantics of Japanese II]. Tokyo: Kurosio Publishers.

Tokieda, Motoki. 1941. *Kokugogaku genron* [Principles of Japanese linguistics]. Tokyo: Iwanami Shoten.

Uehara, Satoshi. 1998. *Syntactic categories in Japanese: A cognitive and typological introduction.* Tokyo: Kurosio Publishers.

Urushibara, Saeko. 1993. *Syntactic categories and extended projections in Japanese.* Waltham, MA: Brandeis University dissertation.

Uchimaru, Yukako. 2006. Dōshi no *te*-kei o tomonau setsu no tōgo kōzō nitsuite [The syntactic structure of clauses with the *te*-form of verbs]. *Nihongo no Kenkyu* 2(1). 1–14.

Washio, Ryuichi. 1997. Resultatives, compositionality and language variation. *Journal of East Asian Linguistics* 6(1). 1–49.

Yamada, Yoshio. 1908. *Nihon bunpōron* [Japanese grammar]. Tokyo: Hōbunkan.

Yamakido, Yoko. 2005. *The nature of adjective inflection in Japanese.* Stony Brook, NY: Stony Brook University dissertation.

Taro Kageyama
14 Lexical integrity and the morphology-syntax interface

1 Introduction

How are morphology and syntax related? Do they interact with each other, and if so, in what way? Are words detached from phrases and sentences, or do they form a continuum? These are among the central issues in morphological theory that have been addressed repeatedly from universal as well as language-particular perspectives but are not settled in any satisfactory way. Presumably due to its agglutinative character, Japanese exhibits a broad variety of phenomena that involve intricate interactions of morphology and syntax. This chapter probes these phenomena to clarify the nature of what has been called "lexical integrity" (or "syntactic atomicity").

Lexical integrity, a cover concept characterizing the linguistic unit of "word" as a syntactically unbreakable chunk of morphological elements, has been formalized in various ways, among which Anderson's (1992) definition is cited below in (1) as an exemplary exposition (see also Lapointe 1980; Selkirk 1982; Di Sciullo and Williams 1987; Bresnan and Mchombo 1995; Di Sciullo 2005; Lieber and Scalise 2006; Booij 2008).

(1)   Principle of lexical integrity
       The syntax neither manipulates nor has access to the internal structure of words.
       (Anderson 1992: 84)

In the 1980s and 1990s, lexical integrity was held to be a universal hallmark of words as strictly distinguished from syntactic phrases, and it played a vital role in the development of theories of "lexicalism" and "syntacticism" in morphology.

As an embodiment of the fundamental distinction between morphology (words) and syntax (phrases and sentences), the principle of lexical integrity is valid in a large measure for many languages including Japanese. Some examples from Japanese are given below that illustrate how the principle works. First, the contrast in (2) demonstrates that a phrasal category (ANP=Adjectival Noun phrase) may appear inside another phrasal category (NP=Noun Phrase) in (2a), but not inside a lexical category (N=Noun) in (2b).

(2) a. Noun Phrase may contain a phrasal category.
         [NP *kookyuu-na*   *hoteru*], [NP [ANP *taihen kookyuu-na*]   *hoteru*]
         [NP high.class-INFL hotel]   [NP [ANP very   high.class-INFL] hotel]

    b. Compound Noun may not contain a phrasal category.

       [_N *kookyuu-hoteru*], *[_N [_{ANP} *taihen kookyuu*]-*hoteru*]

       [_N high.class-hotel] *[_N [_{ANP} very   high.class]-hotel]

This phenomenon is generalized by Roeper and Siegel (1978) and Botha (1983) as the "No Phrase Constraint": Syntactic categories cannot intrude into word structure. Since phrases as well as functional categories like case particles and inflections belong to the realm of syntax, the No Phrase Constraint can be viewed as a partial manifestation of the first half of Anderson's principle of lexical integrity: "The syntax cannot manipulate the internal structure of words".

    Syntactic manipulation refers not only to the embedding of phrasal categories in word structure but also to the movement or deletion of any part of word structure in syntax. The examples in (3) illustrate the difference between a noun phrase and a compound word as to the applicability of "backward gapping", a sentence-level rule deleting a string of words in the first conjunct of a coordinate sentence under identity with the same string of words in the second conjunct (Kageyama 1989, 1993).[1]

(3) a. Part of a noun phrase can be gapped.

     Musuko wa [_{NP} *kokuritu no ~~daigaku~~*] ~~*ni hairi*~~, *(sosite) musume wa*

     son   TOP   national GEN ~~university~~ ~~DAT~~ ~~enter~~ (and)   daughter TOP

     [_{NP} *siritu no* || *daigaku*] *ni*   *hait-ta.*

     private   GEN university DAT enter-PST

     'My son entered a national university, and my daughter a private university.'

    b. Part of a compound word cannot be gapped.

     *Musuko wa [_N *kokuritu-~~daigaku~~*] ~~*ni hairi,*~~ *(sosite) musume wa*

     son   TOP [national-~~university~~] ~~DAT~~ ~~enter~~ (and)   daughter TOP

     [_N *siritu-* || *daigaku*]   *ni*   *hait-ta.*

     [private-   university] DAT enter-PST (same meaning as 3a)

Sentence (3a) contains two noun phrases *kokuritu no daigaku* [national.foundation GEN university] and *siritu no daigaku* [private.foundation GEN university] with the intervening genitive particle *no*, which syntactically separates *daigaku* from *kokuritu/ siritu*; the syntactic independence of the two constituents is reflected in their pronunciation with two accent peaks, as in *koKURITU NO daIGAKU* and *SIritu no daIGAKU* (capital letters stand for high pitches). Sentence (3b), on the other hand, contains two compound nouns without the intervening genitive particle, *kokuritu-daigaku* [national-university] and *siritu-daigaku* [private-university], which are each pronounced

---

1 Kornfilt (2012) regards an analogous phenomenon in Turkish as Right Node Raising, but we main tain that it is a deletion rule rather than movement because two or more elements may be involved as long as they form a phonologically continuous sequence.

as one phonological unit with a single accent peak, as in *koKURITU-DAigaku* and *siRITU-DAigaku*.

Now, what is important about gapped sentences is a characteristic phonological break at the position indicatged by the symbol " || " in the second conjunct – essentially the same kind of phonological break that is observed with English gapping sentences like *John bought a motorcycle, and Mike || a mountain bike* or English right node raising constructions like *This car is similar to, but not exactly identical with, that one*. Thus, (3a) and (3b) are to be pronounced with a short break between *siritu no* and *daigaku* in the former and between *siritu* and *daigaku* in the latter. As shown, the phonological break in (3a) is perfectly normal, but the one in (3b) is abnormal because it breaks up a compound word. Backward gapping thus obeys the first half of Anderson's principle of lexical integrity: The syntax cannot manipulate the internal structure of words.

The effect of the second half of Anderson's principle – "The syntax cannot have access to the internal structure of words" – is illustrated with the example in (4).

(4) An external phrase cannot modify only one constituent of a compound word.
     *taihen   [$_N$ ***kookyuu****-hoteru*]
     very   [$_N$ high.class-hotel]

The example in (4) is ruled out because a degree adverb *taihen* 'very', which occurs outside the compound *kookyuu-hoteru* 'high-class hotel', is intended to modify only the first constituent (*kookyuu* 'high-class') of the compound noun.

As shown above, garden-variety examples like those in (2), (3), and (4) endorse the fundamental correctness of the principle of lexical integrity. More recent studies on a variety of languages, however, have uncovered a number of counterexamples and other problematic phenomena that throw the universal validity of the lexical integrity principle into question. An excellent review of the field is found in Lieber and Scalise (2006).

In this chapter, we will present a diversity of Japanese word formation processes that straddle the boundary of morphology and syntax and are thus challenging to the principle of lexical integrity. Section 2 discusses phenomena where phrasal categories are involved in word structure – genitive compounds (2.1), phrasal compounds (2.2), and Word plus (2.3) – whereas Section 3 examines the phenomena of argument inheritance (3.1) and anaphoric islands (3.2), where syntax appears to have access to a word-internal element. In Section 4, we shift attention to cases where morphological elements show up as part of syntactic structure: clitics (4.1) and Hattori's (1950) *fuzoku-go* or non-independent words (4.2). In Section 5, we clarify the universal and non-universal facets of lexical integrity by disintegrating the principle in (1) into two components. We will also mention the possibility of degrammaticalization of a prefix into an emphatic adverb.

2 Syntactic categories inside words

We will start out with discussion of three kinds of phenomena that appear to involve phrasal categories inside words in apparent violation of the No Phrase Constraint: genitive compounds (Section 2.1), phrasal compounds (2.2), and the novel category of Word plus (2.3). We will show that despite the apparent involvement of phrasal elements, these phenomena can all be subsumed within the realm of lexical word formation.

2.1 Genitive compounds

A reliable diagnosis of "wordhood" in Japanese is to see whether functional categories like case particles, focus particles, and tense inflections can legitimately appear at an internal (not peripheral) position of a given morphological object. Since these grammatical elements function to connect one phrase with another, they are systematically precluded from appearing inside words, as illustrated by (5). This is exactly what the No Phrase Constraint would predict.

(5) a. Case particles are excluded from inside words.
         [N *ara-sagasi*] 'faultfinding', * [N *ara **no** sagasi*] (fault GEN finding),
         *[N *ara **o** sagasi*] (fault ACC finding)

     b. Verbal tense inflections are excluded from inside words.
         [N *aki-ya*] (being.vacant-house) 'vacant house', *[N *ak-**u** ya*]
         (be.vacant-PRS house)

     c. Adjectival tense inflections are excluded from inside words.
         [N *huru-dokei*] (old-clock) 'antique clock', *[N *huru-**i** dokei*] (old-PRS clock)

The asterisked expressions in (5a) and (5b) simply do not count as a word. In (5c), if the voiced [d] due to *rendaku* were changed to voiceless [t], the result would be *huru-i tokei*, which is acceptable only as a noun phrase meaning 'an old clock'.

Despite this general restriction, there are examples where the genitive particle *no* and sometimes the nominative particle *ga* (and very rarely, *na* as in *ma na ko* [eye GEN small.thing] 'pupil of the eye') that used to have a noun-modifying function in classical Japanese occur inside what appear to be compound words. Representative examples are shown in (6).

(6) [N *ko-no-ha*] (tree-GEN-leaf) 'leaf', [N *hi-no-de*] (sun-GEN-rise) 'sunrise',
     [N *kumo-no-su*] (spider-GEN-web) 'spider's web', [N *nomi-no-iti*] (flea-GEN-market)
     'flea market', [N *ama-no-kawa*] (heaven-GEN-river) 'the Milky Way',
     [N *mago-no-te*] (grandchild-GEN-hand) 'back scratcher', [N *ma-no-atari*]
     (eye-GEN-front) '(before) one's very eyes', [N *wa-ga-ya*] (I-GEN-house) 'my home'

English also has similar expressions involving the genitive 's, such as *a spider's web* and *an old people's home*, which Taylor (1996) dubs "genitive compounds". For convenience' sake, the same term will be employed to refer to Japanese examples like those in (6), although it will be made clear shortly that the term "compound" is not quite an accurate description of their morphological nature.

The assumption that genitive compounds like those in (6) count as words can be substantiated by their morphological characteristics, as noted by Kageyama (1999). First, some of them involve bound allomorphs of nouns that are strictly restricted to word-internal position. Thus, *ko* 'tree' in [N *ko-no-ha*] (tree-GEN-leaf), a bound allomorph of the noun *ki* 'tree', is strictly limited to the first position of compounds as in *ko-kage* [tree-shade] 'shade of a tree'. Likewise, *ma* 'eye' in [N *ma-no-atari*] '(before) one's very eyes' is a bound allomorph of the noun *me* 'eye', and in [N *wa-ga-ya*] (I-GEN-home) 'my home', both *wa* 'I' and *ya* 'home' are bound morphemes in contemporary Japanese.

Second, genitive compounds have developed special meanings in contrast with their semantically transparent noun phrase counterparts. For example, [N *mago-no-te*] (grandchild-GEN-hand) designates a back scratcher, as opposed to the corresponding noun phrase [NP *mago no te*], which conveys the literal meaning of 'a grandson's (or granddaughter's) hand'. In some dialects though not in standard Japanese, these two diverge in accentuation as well.

Third, genitive compounds exhibit lexical integrity. In accordance with the No Phrase Constraint, they reject intrusion of syntactic elements like an adjectival modifier, as shown by the contrast between (7a), a genitive compound involving the bound morpheme *ko* 'tree', and (7b), a noun phrase involving a freestanding noun *ki* 'tree'.

(7) a. *[N *ko-no*   [NP ***ooki-i** ha*]] (tree-GEN [big-INFL leaf])   lit. 'tree big-leaves'

   b. [NP *ki no*   [NP ***ooki-i** ha*]] (tree GEN [big-INFL leaf])   'the tree's big leaves'

Fourth, external modifiers are incapable of referring directly to the first elements of genitive compounds. This point is made clear by comparing the genitive compound in (8a) with its corresponding noun phrase in (8b).

(8) a. *ooki-i*   [N ***ko**-no-ha*]   (big [N tree-GEN-leaf])

   b. *ooki-i*   [NP ***ki** no ha*]   (big [NP tree GEN leaf])

In (8a), the adjective *ooki-i* 'big' can never be interpreted as modifying only the bound morpheme *ko* 'tree' in [N *ko-no-ha*]; instead, (8a) can only mean 'big leaves', where the adjective 'big' modifies the whole of [N *ko-no-ha*] 'leaves'. (8b), on the other hand, is ambiguous between the reading 'big leaves' and crucially, the reading 'leaves of a big tree'.

Fifth, genitive compounds may be embedded in a larger compound, as shown in (9a). This sharply contrasts with the total impossibility of embedding a noun phrase within a compound noun, as shown by the ungrammaticality of (9b).

(9) a. [$_N$ [$_N$ *ko-no-ha*] *donburi*] ([$_N$ [$_N$ tree-GEN-leaf] bowl.of.rice])
'a bowl of rice topped with boiled eggs and slices of steamed fish paste'

   b. *[$_N$ [$_{NP}$ *ki no ha*] *donburi*] ([$_N$ [$_{NP}$ tree GEN leaf] bowl.of.rice])

Having shown the word status of genitive compounds, we must raise the question: Do the genitive compounds in themselves count as counterexamples to the No Phrase Constraint? Recall that the No Phrase Constraint outlaws intrusion of a phrasal category into word structure. Genitive compounds do not fall under this characterization because they are not created by embedding a phrasal category in a pre-existing word structure. Rather, they are best considered as resulting from reanalysis of the noun phrase structure 'N *no/ga* N' as a simple noun, as schematically represented in (10).

(10) Reanalysis of a genitive NP as a noun
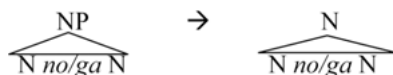
From the standpoint of morphology, 'N-*no/ga*-N' is an unanalyzable whole, although its internal composition is perceptible to native speakers. This is like the English words *would-be* and *must-have*, which result from reanalysis of modal auxiliaries and the verbs *be/have*. Since no compounding process is involved in (10), the term "genitive compound" is, strictly speaking, a misnomer.

There is a universal constraint that the head (i.e. the category-determining constituent) of a complex word cannot be phrasal though non-heads might be (cf. Di Sciullo 2005). This constraint, coupled with the reanalyzed structure in (10), predicts that genitive compounds can be the head of a larger complex word. This prediction is borne out by examples like (11), where [$_N$ *hi-no-de*] 'sunrise' and [$_N$ *nomi-no-iti*] 'flea market' indeed function as the heads (right-hand members) of larger complex words.

(11) a. [$_N$ *hatu*-[$_N$ *hi-no-de*]] (first-[sun-GEN-rise]) 'first sunrise of the year'

   b. [$_N$ *Tookyoo*-[$_N$ *nomi-no-iti*]] (Tokyo-[flea-GEN-market]) 'Tokyo flea market'

We have argued that genitive compounds are irrelevant to the No Phrase Constraint because they are created by an irregular process of reanalyzing a genitive NP as a simple N. Although many of the genitive compounds in current use are historical relics, such category reanalysis still shows some productivity in the creation

of proper names of public interest, such as *Maru-no-uti* [round-GEN-inside] (a place name), *Keiroo-no-hi* 'Senior Citizens' Day', and *Mai-no-umi* [dance-GEN-sea] (name of a *sumo* wrestler).

2.2 Phrasal compounds

In light of the total ungrammaticality of the example given earlier in (2b), *[$_N$ [$_{AP}$ *taihen kookyuu*]-*hoteru*] ([very high.class]-hotel), we must acknowledge that the No Phrase Constraint (Botha 1983) has a certain validity to it. Nevertheless, a number of counterexamples under the rubric of "phrasal compounds" have been reported from a variety of languages (Lieber 1992 for English; Seki 1977; Kageyama 1993 for Japanese; Booij 2010 for Dutch). Based on English phrasal compounds like *a pipe-and-slipper husband*, *over-the-fence gossip*, and *a slept-all-day look*, Lieber (1992) went so far as to remove the boundary between words and phrases altogether, proposing that all complex words are produced by syntactic rules, although the same author has now relinquished that idea in Lieber and Scalise (2006).

Just because many languages have long compounds whose non-heads are phrasal or even sentential categories does not mean that the boundary between words (morphology) and phrases (syntax) should be eliminated altogether. On the contrary, it is possible to bring order out of such disorderly compounds if we can properly sort them into coherent groups on the basis of the lexical nature of the head and non-head elements. Specifically, phrasal compounds whose non-head constituents are regarded as fixed idioms, as in *pipe and slippers* and *over the fence*, will make up a class. Idiomatic phrases can be naturally reanalyzed as simplex words and become susceptible to embedding in larger compounds. Another class has to do with those phrasal compounds whose heads are limited to lexically designated nouns like *look* and *attitude*. These particular head nouns allow their non-head elements to be expanded to phrasal categories. It appears that Japanese phrasal compounds also fall into these two classes.

Japanese compounds that have more or less lexicalized phrases in their non-head position are illustrated by actually occurring examples like those in (12) (a, b, c from Kageyama 1993; d from Martin 1975: 879).

(12) a. [[***hi-no-mi***] *yagura*] ([fire-GEN-watch] tower) 'a fire lookout'

 b. [[***aka-i hane***] *bokin*] ([red-PRS feather] fund.raising) 'a Red Feather drive for charity'

 c. [[***toretore no hugu***]-*ryoori*] ([fresh.caught GEN globefish]-dish) 'dishes of fresh-caught globefish'

 d. *Sekai no Kurosawa mo* [[***kako no hito***]-*si*]-*sare* (world GEN Kurosawa even [past GEN person]-regard-PASS) 'the world-famous Kurosawa (Japanese movie director) is regarded as a has-been'

(12a) embeds a genitive compound in the non-head position, and (12b) a fixed A+N combination *aka-i hane* 'red feather'. In (12c), although the whole non-head phrase 'fresh-caught globefish' cannot be regarded as a genitive compound or a lexicalized expression, *toretore no* 'fresh-caught' is an idiomatic phrase depicting the freshness of raw fish. In fact, it cannot be replaced by a purely descriptive adjective like 'big' as in **[ooki-i hugu]-ryoori* ([big-INFL globefish]-dish) 'dishes of big globefish'. In (12d), *kako no hito* 'a has-been' is also a fixed expression.

More common and more productive than this class is the other class of phrasal compounds whose heads are lexically designated. Some attested examples are selected in (13), where the non-head constituents enclosed in square brackets are neither fixed nor idiomatic phrases and can be replaced freely and productively by other phrasal expressions. For more examples, the reader is referred to Kageyama (1993).

(13) a. [*izime no nai akaru-i gakkoo*]-***zukuri*** ([bullying GEN nonexistent cheerful school]-making) 'making a bully-less, cheerful school'

b. [*hiro-i niwa*]-***tuki*** ([large-INFL garden]-be.equipped) 'equipped with a large garden'

c. [*sekai no neko*]-***ten*** ([world GEN cat]-exhibition) 'exhibition of cats of the world'

d. [*goruhukai no daiitininsya*]-***teki*** ([golf.world GEN leader]-like) 'like the leader in the golf world'

e. [*kani-ryoori to onsen*]-***koosu*** ([crab-dishes and hot spas]-course) 'tour for crab dishes and hot spas'

f. [*karaoke to geemu*]-***taikai*** ([karaoke and game]-tournament) 'tournament of karaoke and computer games'

In (13), the morphological composition indicated by square brackets is overshadowed by the phonological grouping by which the boldfaced elements are pronounced as an uninterrupted sequence with a single accentual peak. The phonological grouping, however, should not be taken to indicate that the elements that are not boldfaced in (13) are syntactically or morphologically separated from the boldfaced elements. On the contrary, the whole sequences qualify as words as they can be embedded in the non-head position of a larger word, as in (14a).

(14) a. [[*hiro-i   niwa-tuki*]   [*kookyuu-zyuutaku*]]
[[large-INFL garden-be.attached] [luxury-residence]
'a large-gardened luxury residence'   cf. (13b)

b. **kookyuu* [[*hiro-i   niwa-tuki*]   *zyuutaku*]
   luxury   [[large-INFL garden-be.attaached] residence]
   'a luxury residence with a large garden'

c. *kookyuu* [*niwa-tuki   zyuutaku*]
   luxury   [garden-be.attached   residence]
   'a luxury residence with a garden'

Particularly telling is the asymmetry between the grammatical compound in (14a), where the phrasal compound [*hiroi niwa-tuki*] 'with a large garden' is expanded to the right, and the ungrammatical compound in (14b), where the same compound is expanded to the left. (14b) is ruled out because it contains a phrasal compound [[*hiroi niwa-tuki*] *zyuutaku*] in its head position. The ungrammaticality of (14b) is ameliorated only by removing the inflected adjective *hiro-i* 'large' from it, as in (14c). The contrast between (14a) and (14b) thus indicates that a phrasal compound can appear only in a non-head position, but not in (part of) the head position. This means that the phrasal categories embedded in phrasal compounds are "visible" to morphology.

The preceding observations demonstrate that phrasal compounds count as true exceptions to the No Phrase Constraint. What is notable about all these Japanese phrasal compounds is that their heads appear to be limited to designated lexemes, whether compounding elements or suffixes. The heads *-zukuri* 'making' and *-tuki* '(be accompanied) with' in (13a) and (13b), which are nouns converted from native verbs (*tukuru* 'make', *tuku* 'attach'), are limited to the second position of compounds, so that they are plausibly characterized as suffix-like. There are quite a few deverbal nouns of the native Japanese stratum that behave similarly, such as *-muke* 'intended for', *-atukai* 'treat as', *-iri* 'containing', and *-ate* 'addressed to'. Many Sino-Japanese suffixes like *-ten* 'exhibition' (13c) and *-teki* 'like' (13d) also accept phrasal bases, as do *-kyuu* 'class', *-taku* 'house', *-huu* 'style', and *-yoo* 'designed for'. Strictly speaking, these examples should be termed "phrasal derivatives", but we will continue to use the term "phrasal compounds". On the other hand, the examples in (13e) and (13f) are headed by content words rather than suffix-like elements. Their usage, however, seems limited to catchphrases in fliers for commercial advertisement.

To summarize so far, phrasal compounds are observed fairly frequently in written and spoken language of casual style, but their productivity is constrained both by the head elements and by the conventionalized non-head expressions. In this connection, the behavior of the head noun *kai* 'association' is of some interest. This noun selects a clausal complement headed by a verb that has the present tense inflection, as illustrated by the name of a hypothetical association in (15).

(15) [*Nihon no   onsen   o   mamor-u*]   *kai*
     [Japan GEN hot.spa ACC preserve-PRS] association
     'The Association for Preserving Japanese Hot Spas'

This construction is peculiar if it is analyzed as a compound. First, it contains a tensed verb *mamor-u* [preserve-PRS], despite the fact that tense inflections are generally excluded from inside compounds. Second, unlike the phrasal compounds in (13) which are rather limited in productivity, the *kai* construction of (15) has absolutely no limit to its productivity and any new association or group can be named using this pattern. These peculiarities suggest that (15) might be better viewed as a noun phrase than as a compound. Notice, however, that the tense on the verb is limited to the present tense denoting a habitual or permanent activity rather than an actual activity that takes place at a particular time. More importantly, the construction of (15) can be enlarged to a larger compound by adding another head noun to its right, as in [[[*Nihon no onsen o mamor-u*] *kai*] *kaityoo*] ([[[Japan GEN hot.spa ACC preserve-PRS] association] president]) 'president of the Association for Preserving Japanese Hot Spas'. These semantic and morphological properties will lead us to identify the *kai* construction of (15) as a phrasal compound.

The nature of Japanese phrasal compounds can then be captured by saying that they are basically licensed by particular suffix-like elements in the head position, which will be characterized as "stretchable suffixes" – stretchable in the sense that their bases (non-head elements to which they attach) can be stretched from arche-typical categories of bound morphemes to freestanding words to phrasal or senten-tial categories. In fact, all these stretchable suffixes have the versatility of attaching to a variety of categories ranging from bound roots and words to phrases, as shown in Table 1.

Table 1: Stretchable suffixes

	Bound root as base	**Word as base**	**Phrase as base**
huu 'style' (S-J)	*yoo huu* 'Western-style'	*Huransu huu* 'French-style'	*[tyuusei no Huransu] huu* 'in the style of medieval France'
ten 'exhibition' (S-J)	*ko ten* 'private exhibition'	*syasin ten* 'photo exhibition'	*[syuururearizumu no kyosyootati] ten* 'exhibition of works by giants of surrealism'
tuki 'be equipped with' (native)	DNA	*niwa tuki* 'with a garden'	*[hiroi niwa] tuki* 'with a large garden'

Attachment to phrasal categories, however, must be deemed an exceptional behavior of suffixes because not all suffixes allow this possibility freely. We conclude that phrasal compounds, though ostensibly deviant from the No Phrase Constraint, are manageable counterexamples and do not drastically undermine the basic tenet of the morphology-syntax separation.

2.3 Word plus

This subsection gives an overview of a special morphological category "Word plus" (W+), advocated by Kageyama (1993, 2001). Word plus looks like a phrasal category in terms of its phonological behavior, and yet it belongs to the realm of morphology. The basic idea for this novel category derives from a peculiar set of phrase-like prefixes of the Sino-Japanese stratum that exhibit their own accent instead of following the regular lexical accent, as observed by Aoyagi (1969), Kageyama (1982), and Poser (1990). Representative examples are shown in (16).

(16)   a.   *hon-* 'this, the present': *HOn | daigaku* 'this university'

  b.   *moto-* 'former, ex-': *MOto | sooridaizin* 'a former Prime Minister'

  c.   *zen-* 'ex-, immediately preceding': *ZEn | sooridaizin* 'the ex-Prime Minister'

  d.   *gen-* 'current': *GEn | kaityoo* 'the current president'

  e.   *kaku-* 'each': *KAku | daigaku* 'each university'

  f.   *boo-* 'a certain': *BOo | tyomeisakka* 'a certain famous writer'

  g.   *doo-* 'above-mentioned': *DOo | sya* 'the above mentioned company'

  h.   *zen-* 'all, whole': *ZEn | kumiaiin* 'all the union members'

  i.   *ryoo-* 'both': *RYOo | si* 'both cities'

  j.   *ko-* 'deceased': *KO | uemura-si* 'the late Mr. Uemura'

  k.   *han-* 'anti-': *HAn | taisei* 'anti-establishment'

  l.   *hi-* 'non-' : *HI | yooroppa-gengo* 'non-European languages'

In each example of (16), capital letters indicate a high-pitched mora and the symbol " | " stands for a minor phonological break characteristically observed with these prefixes. Thus, the first example in (16a), *HOn | daigaku* 'this university', is pronounced with a high pitch on the initial mora *HO*, followed by a low pitch on the following [n], and is then accompanied with a very short break much like the one observed after *non-* in English words like *non-European*; the base words after the prefixes are pronounced with low pitches.

Because of their semantic function to delimit the following nouns in some way or other, these prefixes are sometimes called "determiner-like (or quantifier-like) prefixes" (*rentaishi-teki* 'noun-modifier-like'). However, meaning cannot be a definitive criterion for this category. For example, the prefix *soo-* 'all, total', which adds a quantificational meaning to the base noun, as in *soO-SYUunyuu* 'a total income', is not a Word plus because it is pronounced with the regular lexical accent. The phrase-like accentuation as shown in (16) is thus the defining feature of W+.

Some of the phrase-like prefixes are polysemous, showing both the lexical accent pattern and the phrase-like accentuation. One such prefix is *doo-* 'same', as in (17).

(17) a. Lexical accent: *doO-ZIkoku* 同時刻
   same-time   'the same (identical) time'

   b. Phrasal accent: *DOo | zikoku*   同時刻
   aforementioned | time   'the time mentioned earlier'

The two examples in (17), albeit identical when written in Chinese characters, differ in accentual pattern and semantic interpretation. The lexical accent contour observed in (17a), one that is regularly observed with compound and derived words of all lexical strata, serves to mark the word unit by consolidating the first and second members with a single stretch of accent peak. Thus, the continuous high pitches on "O" and "ZI' in *doO-ZIkoku* function to bridge the morphological boundary between the two members. By contrast, the bridging function of a high-pitched sequence is not observed in (17b), where the prefix displays its own accent, followed by a brief phonological break. (17b) thus lacks the phonological solidarity that is normally expected of word units. Nonetheless, there is good reason to identify the prefixes at issue as morphological rather than syntactic entities, contrary to Miller (1993), who regards them as syntactic determiners that are attached to base nouns by movement in syntactic structure. Among the tests Kageyama (2001) exploited to show the word status of complex expressions involving these prefixes, the most straightforward one is the exclusion of phrasal categories.

(18) a. *[*boo* | [_NP ***yuumei-na*** *haiyuu*]]
        [certain | [famous-INFL   actor]]
        'a certain famous actor'

   b. *[*kaku* | [_NP ***tihoo***   ***no***   *tosi*]]
        [each | [province GEN city]]
        'each provincial city'

The expressions in (18a) and (18b) are ungrammatical because of the violation of the No Phrase Constraint. As expected, the ungrammaticality will disappear if the ending *na* and the genitive particle *no* inside the noun phrases are removed in such a way that *yuumei-haiyuu* 'famous actor' and *tihoo-tosi* 'provincial city' are realized as compound nouns.

Kageyama (2001) further observed that the phrase-like accentuation with a minor phonological break is not a prerogative of the particular set of Sino-Japanese prefixes as assumed in previous literature (Aoyagi 1969; Poser 1990), but is actually shared by long Sino-Japanese compounds like those in (19).

(19) a. [*koKURITU-DAigaku* | *gakutyoo*]
    [national-university | president]
    'president of a national university'

   b. [*ZEnkoku* | *gassyoo-konkuuru*]
    [all-Japan | chorus-contest]
    'All-Japan chorus contest'

These long compounds are also equipped with lexical integrity, as demonstrated by the unacceptability of (20), where the genitive particle *no* intervenes in compounds.

(20) a. *[*kokuritu no daigaku* | *gakutyoo*]
    [national GEN university | president] Cf. (19a)

   b. *[*zenkoku* | *gassyoo no konkuuru*]
    [all-Japan | chorus GEN contest]   Cf. (19b)

The minor phonological break shared by the S-J prefixed words and the S-J compound words indicates that these two sets of words belong to the same category. This suggestion is buttressed by the recursivity of W⁺ categories, by which compounds and derived words at issue can be embedded inside each other to produce more and more complex words with iterated minor breaks, as illustrated in (21).

(21) a. [[*kokuritu-daigaku*] | [*zen* | *gakutyoo*]]
    [[national-university] | [former | president]]
    'former president of a national university'

   b. [*boo* | [[*kokuritu-daigaku*] | [*zen* | *gakutyoo*]]]
    [certain | [national-university] | [former | president]]]
    'a certain ex-president of a national university'

Based on these and many other observations, Kageyama (2001) proposes Word plus (W⁺) as a morphological category that is larger than the ordinary Word but is still not syntactic. While Lieber and Scalise (2006) suggest that the W⁺ phenomena might be analyzed as prefixation or compounding on a phrasal basis, the exclusion of phrasal categories as shown in (18) and (20) and the recursive application shown in (21) indicate that they are genuinely morphological categories. The relationship between various categories of morphology and syntax can then be organized as in (22).

(22)

| syntactic categories | { XP, X′ | } phrasal accent |

| morphological categories | Word plus, { Word, Root | } lexical accent |

(22) shows that morphological structure is composed of three categories: Root (bound morpheme), Word, and Word plus. Although we cannot go deeply into the distinction of Root and Word here, a fairly reliable diagnosis is the possibility of embedding certain coordinate conjunctions. As observed by Kageyama (1982, 1993), Japanese exploits several coordinate conjunctions that can be differentiated by the morphological categories they can conjoin. The conjunctions *naisi* 'or' and *oyobi* 'and' are capable of conjoining compound words and W⁺ words, but not bound morphemes, whereas the conjunction *to* 'and' applies to noun phrases and (rather marginally) to W⁺, but not to compound words or bound roots. Table 2 gives a rough idea of how these three conjunctions work in morphological structure.

Table 2: Coordinators inside words

	naisi 'and/or', *oyobi* 'and'	*to* 'and'
Inside Root category	*[*koo **naisi/oyobi** ryoku*]*-tya* [black or/and green]-tea	*[*koo **to** ryoku*]*-tya* [black and green]-tea
Inside Word category	[*kooritu **naisi/oyobi** siritu*]*-daigaku* [public or/and private]-university	*[*kooritu **to** siritu*]*-daigaku* [public and private]-university
Inside Word⁺ category	*gen* | [*kaityoo **naisi/oyobi** hukukaityoo*] current [president or/and vice.president]	?*gen* | [*kaityoo **to** hukukaityoo*] current [president and vice.president]

Notice that the conjunctions *naisi* and *oyobi* appearing inside compound words are irrelevant to the No Phrase Constraint because they are morphological objects in themselves.

As is frequently pointed out (see Chapter 1 [Kageyama and Saito, this volume]), Sino-Japanese words, as compared with native Japanese words, enjoy full productivity in word formation, producing infinitely long words. The source of this infinite productivity is now accurately pinpointed to the special category W⁺.

3 Syntactic access to word-internal structure

This section discusses two kinds of phenomena in which syntax appears to look into the internal structure of a word: inheritance of arguments from non-head constituents (Section 3.1), and anaphoric islands and word-internal *wh*-words (3.2). These phenomena present a challenging problem for the second half of the principle of lexical integrity given at the outset in (1): "The syntax cannot have access to the internal structure of words".

3.1 Argument inheritance from non-heads

The "no syntactic access" part of the lexical integrity principle properly rules out cases where an external modifier refers to a part of a word, as shown by the ungram-

maticality of **taihen* [_N *kookyuu-hoteru*] (very [high.class-hotel]) given in (2b). This constraint holds generally for cases in which external modifiers are adverbials or adjuncts. If an external phrase represents the argument of a predicate member of a complex predicate, however, it is capable of accessing the word-internal predicate constituent. This situation, generally called "argument inheritance" or simply "inheritance", is divided into two types.

In one type, the categorial head of a compound predicate feeds argument information to syntax.

(23) a. Inheritance from heads
         mondai *no*   [*sooki-**kaiketu***]   (Sugioka 1989)
         problem GEN [early-resolution]
         'an early resolution of the problem'

     b. ***take***   *o*   *mapputatu ni*   [*tataki-**waru***] (Kageyama 1996)
         bamboo ACC right.half   DAT [strike-split]
         'to split a bamboo in half by striking it'

(23a) exemplifies a compound VN of the form "N-VN", and (23b) a compound verb of the form "V-V". In both cases, the external expressions ('problem' and 'bamboo') are associated with the right-hand constituents of the compounds, *kaiketu* 'resolution' and *waru* 'split' in such a way that the 'problem' corresponds to the object argument of 'resolution' and the 'bamboo' to the object of 'split'. Inheritance of argument structure information from the head elements is a universally attested phenomenon (Di Sciullo and Williams 1987; Roeper 1987). It is important to stress that inheritance does not count as a violation of the principle of lexical integrity. This is so because the external phrases ('problem' and 'bamboo') in (23) are not peeping through the bracketed compound structures to directly refer to the head constituents; instead, the argument structure information of the heads is inherited (i.e. passed on) to the whole compounds and then is projected to the external phrases. Inheritance is thus a legitimate way to evade the effects of the "no syntactic access" constraint.

In contradistinction to inheritance from heads, inheritance from non-head constituents on the left-hand side is explicitly disallowed for English by Roeper (1987). This restriction, however, is not universal because Japanese allows this type of inheritance systematically. Observe the compound words in (24) (examples of this kind are also taken up in Chapter 6 [Namiki and Kageyama, this volume]).

(24) Inheritance from non-head constituents
     a. <u>*kozutumi*</u> *no*   [***haitatu**-nin*]   (Sugioka 1989)
         package GEN [delivery-man]
         lit. 'the delivery man of the package'

     b. *Titioya wa* <u>*musuko*</u> *o*   [***sikari**-tuke*]-*ta.*   (Yumoto 2005)
         father   TOP son   ACC [scold-do.violently]-PST
         'Father scolded his son thoroughly.'

In these examples, the underlined external nouns are associated with the boldfaced non-head constituents of the compounds, so that the 'package' is interpreted as the object argument of 'delivery', and the 'son' as that of 'scold'.

Kageyama (1993) further observes cases where inheritance of argument structure information takes place from both head and non-head constituents concurrently, as illustrated in (25).

(25) ?*Watasitati wa natu   no iti-ya   o   gakusei-zidai no omoide o*
     we   TOP summer GEN one-night ACC student-days GEN memory ACC
     [*katari-akasi*]-*ta*.
     [tell-pass]-PST
     'We talked the night away about the memory of our school days.'

The example of (25), albeit slightly awkward because of the simultaneous occurrence of two accusative phrases, involves double inheritance, whereby the temporal path phrase *iti-ya o* 'one night' is linked with the head verb *akasi-* 'pass (a night)' in the compound verb (i.e. *iti-ya o akasu* [one-night ACC pass] 'pass a night'), and the theme phrase *omoide o* 'the memory' with the non-head verb *katari-* 'talk' (i.e. *omoide o kataru* [memory ACC talk] 'talk about the memory'). Based on such cases, Kageyama (1993) proposed, in line with Selkirk (1982), Di Sciullo and Williams (1987), and others, that the inheritance principle should be relativized so that argument structure information that is not contained in the head may be inherited from the non-head constituent.

The source of argument structure for inheritance is not confined to predicates like verbs and VNs; it is also found in a certain class of nouns. Kageyama (1993) brings to light examples like those in (26).

(26) a. <u>*sakunen sinda sohu*</u> *no* [**haka**-*mairi*] (last.year died grandfather GEN [grave-visit]) 'a visit to the grave of my grandfather, who died last year'

      b. <u>*rinzin*</u> *no* [**ara**-*sagasi*] (neighbor GEN [fault-finding]) 'finding the fault of neighbors'

      c. <u>*dan-zyo*</u> *no* [**en**-*musubi*] (man-woman [match-making]) 'matchmaking for men and women'

In each example of (26), the bracketed elements are commonly considered to make up a compound word of the form "N + deverbal N", and the underlined external expressions are semantically associated with the boldfaced nouns that appear in the non-head positions of the compounds. Thus in (26a), *sakunen sinda sohu* 'grandfather who died last year' is understood to refer to the person who is buried under the grave denoted by the noun *haka* of the compound *haka-mairi* [grave-visit], so

that the whole example of (26a) means 'a visit to the grave of my grandfather, who died last year'. The same manner of semantic interpretation applies to (26b) and (26c) as well. As Kageyama (1993) showed, such expressions cannot be analyzed as phrasal compounds of the kind discussed in Section 2.2 because adjuncts may intervene between the bracketed parts and the preceding genitive modifiers. If the external nouns were assumed to directly modify the boldfaced nouns across the boundary of compound words, then, these examples would be an egregious violation of lexical integrity.

According to Kageyama (1993), the inheritance from non-head constituents in examples like those in (26) is made possible by a special property of the boldfaced nouns in the non-head position. Crucially, nouns like *haka* 'grave', *ara* 'fault', and *en* 'match' differ from ordinary nouns like 'stone' or 'car' in that their meanings cannot be defined in isolation without reference to other entities that have an intrinsic relation to them. Compare the noun *haka* 'grave' in (26a) with the noun *haka-isi* 'tombstone', for instance. Although these two nouns are similar in meaning, they differ on an important point. Because a tombstone is merely a physical entity made of stone for the purpose of burial, if we add a possessive phrase as in *sohu no haka-isi* 'my grandfather's tombstone', 'my grandfather' is most likely understood as the owner of the tombstone. On the other hand, if *sohu no* is added to the noun *haka* 'grave', the most plausible interpretation will be that my grandfather, a dead person, is buried under the grave. The latter interpretation is made possible by the lexical semantic representation of the noun *haka* 'grave' as 'THE PLACE UNDER WHICH x IS BURIED', where the variable x must be specified externally by a noun representing a particular person. In other words, *haka* is a kind of complement-taking noun discussed in Chapter 18 (Nishiyama, this volume). Given this, the modification relation between *sohu* 'grandfather' and *haka* 'grave' in (26a) is licensed by linking the external noun *sohu* 'grandfather' to the variable x specified in the lexical semantic representation.

This analysis predicts that the external complement of *haka* 'grave' cannot designate anything other than the person who is buried under it, and that the noun *haka-isi* 'tombstone' cannot participate in the argument inheritance phenomenon under discussion. These predictions are both borne out by the ungrammaticality of (27).

(27) a. **sakunen sinda sohu   no   [haka.isi-migaki]*
       last.year died   grandfather GEN   tombstone-polishing
       Intended: 'polishing the tombstone of the grave of my grandfather,
       who died last year'

    b. **dairiseki no   [haka-mairi]*
       marble   GEN   [grave-visit]
       Lit. 'a visit to a grave made of marble'

The nouns *ara* 'fault' (26b) and *en* 'match' (26c) are also considered to have a variable position in their lexical semantic representations, as in 'INHERENT WEAKNESS OF x' for *ara* 'fault' and 'AN INTIMATE TIE BETWEEN x AND y' for *en* 'match'.

To sum up, the phenomenon of external modification in (26) is not actually a violation of lexical integrity because the modification relation is lexically licensed by the variable contained in the semantic representation of the left-hand nouns in compound words.

3.2 Anaphoric islands and word-internal interrogatives

It was Postal (1969) who first pointed out that sentence-level anaphora cannot look into the internal structure of words. The "Anaphoric Island Constraint" is now looked upon as a manifestation of the "no syntactic access" constraint of lexical integrity. The implementation of Postal's (1969) constraint is illustrated by the contrasts between (28a) and (28b), and between (29a) and (29b).

(28)   a.   Mary wants to open a shop in New York because she is from there.

      b. *Mary wants to open a shop in **New York** because she is a **there**-*er*.
         [inbound anaphora]

(29)   a.   Mary is from New York and plans to open a shop there.
         (*there* = New York)

      b. #Mary is a **New York**er and plans to open a shop **there**.
         [outbound anaphora]

Subsequent research showed that neither prohibition of inbound anaphora as in (28b) nor prohibition of outbound anaphora as in (29b) is entirely correct. On the one hand, Sproat (1988) argued that the impossibility of inbound anaphora falls out from the categorial incompatibility of pronouns like *he* and *it*, which are DPs, with word structure, which is necessarily X^0 or lower. On the other hand, Ward, Sproat, and McKoon (1991) demonstrated that outbound anaphora is possible if the referent is properly evoked by inference in a pragmatic context, as in *Officials in the Danish capital believe they've found a way to stop [**bicycle** thefts] – let people use **them** for free*. Yamanashi (1992) presents similar discussion on pragmatically motivated outbound anaphora in Japanese.

Compared with ample discussions on the role of pragmatics in outbound anaphora, little attention has been paid to the nature of inbound anaphora. Here I show two sets of phenomena related to inbound anaphora in Japanese. One concerns compound words involving deictic pronouns, as exemplified in (30).

(30) a. A wife talking about her husband:
    [***kare**-gonomi*] *no   nekutai*
    [he-be.fond]   GEN necktie
    'a necktie of his taste'

   b. Referring to a train arriving at the terminal:
    Kono   densya   wa   [***koko**-domari*]   *da.*
    this   train   TOP [here-stop]   COP
    'This is the end of the line.'

In (30), *kare* 'he' and *koko* 'here' show up inside compounds. These words, however, are deictic pronouns that directly point to particular entities rather than referential pronouns (pronominals) that require linguistic antecedents in discourse. In fact, it is also possible to replace such deictic pronouns with proper names like *Suzuki-sensei* 'Prof. Suzuki' and *Nishinomiya(-eki)* 'Nishinomiya (station)', as in [*Suzuki-sensei-gonomi*] 'suiting the taste of Prof. Suzuki" and *Nishinomiya-domari* 'terminating at Nishinomiya'. The examples in (30) are thus irrelevant to Postal's (1969) Anaphoric Island Constraint, a condition on referential pronouns.

On the other hand, a real case of a word-internal pronominal is found with complex words involving the W$^+$ prefix *doo-* 'aforementioned', pronounced as [DOo] with a high pitch on the initial mora and a short break after it (see Section 2.3). Kageyama (2001) provides ample demonstration that this prefix has the function of creating word-internal pronominals (referential nouns). Observe the following examples.

(31) a. *Daitooryoo   wa   asu   yuukoo-zyooyaku$_i$ ni   tyooin-suru.*
    president   TOP tomorrow amity-treaty$_i$ DAT sign-do

    [***Doo*** | ***zyooyaku$_i$*** | *saisyuu-an*]   *niyoruto* …   (inbound anaphora)
    [said | treaty$_i$ | final-version] according.to

    'The President is going to sign the amity treaty. According to the final version of the said treaty, …'

   b. *Kodai   no   iseki   ga   [**Nara-zyosi-daigaku$_i$** | koonai]*
    ancient GEN ruins   NOM [Nara-women's-university | campus]

    de   mitukat-ta
    LOC be.found-PST

    *to,   **doo** | **daigaku$_i$** no   hakkutu-tyoosa-iinkai*
    COMP said | university GEN excavation-research-committee

    ga   happyoo-si-ta.
    NOM announce-do-PST

    '"Ancient ruins were discovered on the campus of Nara Women's University," the excavation committee of the said university announced.'

In (31), the prefixed words *doo | zyooyaku* 'the said treaty' and *doo | daigaku* 'the said university', embedded in larger compound words, refer to the co-indexed NPs given in the same sentence or discourse in which they occur. This phenomenon is observed with both inbound anaphora (31a) and outbound anaphora (31b). What is unique about such *doo*-prefixed words is that, unlike the deictic pronouns in (30), they cannot directly point to concrete entities given in context. It is thus totally impossible to utter **Watasi wa doo|daigaku ga suki da* 'I like the said university' out of the blue. Rather, the *doo*-prefixed words must refer to linguistic antecedents that appear in well-defined syntactic environments constrained by Condition B of the Binding Theory (see Kageyama (2001) for more details on this point). Consequently, the *doo*-prefixed words count as true counterexamples to the Anaphoric Island Constraint.

Apart from the *doo*-prefixed words, I know of no morpheme of the native Japanese stratum that allows participation in discourse anaphora. The special referential function of *doo*-prefixed words can no doubt be attributed to the special status of the W+ category, which lies at the intersection of words and phrases/sentences.

At this juncture, it will be appropriate to mention an analogous phenomenon in which the syntax appears to have access to word-internal elements – this time, interrogative pronouns. Di Sciullo and Williams (1987) argue that the ungrammaticality of English sentences like (32) constitutes evidence for lexical integrity.

(32)　*The [who-killer] did the police catch?

The word-for-word translation of (32) in Japanese, however, turns out to be grammatically acceptable, as shown in (33a), and even more complex word-internal interrogatives are possible, as illustrated in (33b, c) (Kageyama 1993).

(33) a. *Keisatu wa* [***dare**-gorosi*] *o　tukamae-ta no?*
　　　　 police　TOP [who-killer]　ACC catch-PST　Q　　Cf. (32)

　　 b. *Rinkaan wa　Amerika no　*[*dai.**nan**-dai-daitooryoo*] *desu ka?*
　　　　 Lincoln　TOP America　GEN [what-number-president] COP　Q
　　　　 'What number president of America was Lincoln?'

　　 c. *Anata wa* [***nani**-daigaku | **nani**-gakubu*]　　*o　sotugyoo-si-masi-ta　ka?*
　　　　 you　　TOP [what-university | what-department] ACC graduate-do-POL PST Q
　　　　 'You graduated from which department of what university?'

The Japanese pronouns *dare* 'who' and *nani* 'what' are, strictly speaking, not dedicated interrogatives but are "indeterminates" (Martin 1975; Nishigauchi 1990), which may have a variety of functions depending on the syntactic constructions in which they are used, as in *Dare mo ko-nai* [who also come-NEG] 'Nobody comes' and *Dare demo ko-reru* [who even come-POTEN] 'Anyone can come'. What is striking

about the examples of (33) is that the indeterminate pronouns embedded in complex words are associated with the question particles (*no* or *ka*) appearing at the end of the sentences. In other words, the force of the interrogative sentence final particle penetrates across the word boundary to refer to the indeterminate pronouns inside the complex words. Based on the assumption that Japanese has no overt *wh*-movement in syntax, Kageyama (1993) extends Nishigauchi's (1990) analysis of indeterminate pronouns appearing inside noun phrases to word structure. On this analysis, the [+indeterminate] feature of *dare* 'who' and *nani* 'what' is inherited to the entire complex word, and the question particles *no* and *ka* are linked to this inherited [+indeterminate] feature rather than directly referring to the word-internal indeterminates. The inheritance approach, also utilized in Section 3.1 for inheritance from non-head constituents, can thus evade the violation of the lexical integrity principle.

4 Morphological elements in syntactic structure

In contrast to the preceding two sections dealing with phenomena in which phrasal or phrase-like elements participate in lexical word formation, the present section will tackle reverse cases where morphological elements perform a vital function in the formation of syntactic constructions. A typical example is cliticization, whereby a clitic – a bound morpheme that has an independent status in syntactic structure – is phonologically attached to its host. In addition to clitics, Japanese attests to the existence of rather rare cases where a morphologically bound morpheme keeps its syntactic status without being cliticized to its host. Hattori's (1950) discussion directly pertains to the distinction between cliticized and non-cliticized bound morphemes in syntax.

Hattori (1950) proposed to differentiate what he called *fuzoku-go* 附属語 (translated here as "non-independent words": Vance 1993) from *fuzoku-keishiki* 附属形式 (translated here as "bound affixes") in terms of syntactic independence. Bound affixes, as commonly assumed, are morphologically glued to their bases and are always pronounced as a phonological unit with the bases. On the other hand, non-independent words, though morphologically bound, occupy their own position in syntactic structure without phonologically adjoining to any surrounding word. According to Hattori (1950), the difference can be demonstrated by testing whether a given morpheme can be separated from its base or host by means of a pause or some linguistic material. For example, the suffix *-teki* in *Nihon-teki* [Japan-SUF] 'characteristically Japanese' is a bound affix (*fuzoku-keishiki*) because it can never be uttered separately from its base, as shown by the bizarreness of **Nihon | teki* [Japan | SUF] or **Nihon dake teki* [Japan only SUF]. All the affixes discussed in Section 2 are bound affixes in this sense. Case particles, by contrast, can be

separated from their host, as in *Nihon | ga* [Japan | NOM] or *Nihon **dake** ga* [Japan only NOM]. Case particles as well as other kinds of particles like focus particles and sentence final particles are thus categorized as *fuzoku-go*. Because Hattori (1950) himself was not sufficiently explicit about the relation between clitics and *fuzoku-go*, however, the notion of *fuzoku-go* has tended to be regarded as identical to that of "clitics" in the literature. In this section, we will show that *fuzoku-go* or non-independent words, which maintain their syntactic position even though they are morphologically bound, should be differentiated from clitics, which have a syntactic status but are phonologically attached to their hosts.

4.1 Syntactic suffixes as clitics

It has long been noticed that nominalization often applies to syntactic clauses. A case in point is the adjective-nominalizing suffix *-sa*, which not only produces lexical nouns by attaching to simple adjectives and ANs (*utukusi-sa* 'beauty', *ganko-sa* 'stubbornness') but, more importantly for the purpose of the present chapter, is capable of adjoining to *hosi(i)* 'want to have' and other complex predicates that are created in syntactic structure (Kageyama 1982; Sugioka 1986; Chapter 10 [Sugioka and Ito, this volume]). The syntactic nature of the latter usage of *-sa*, manifests itself in (34).

(34) a. *Otoko wa* [*kane   ga   **hosi**]*-**sa***   *ni   ginkoo-gootoo o   hatarai-ta.*
man   TOP [money NOM want]-NMLZ DAT bank-robbery ACC commit-PST
'The man robbed a bank out of a desire to get money.'

b. [*kodomo ni   oisii   mono o   **tabe-sase-te yari-ta**]*-**sa***   *ni*
[kid   DAT yummy food   ACC eat-CAUS GER give.favor-want]-NMLZ DAT
'from a desire to make his kids eat good food'

Syntactically, the suffix *-sa* takes the bracketed clauses in its scope, whereas phonologically, the boldfaced sequences *hosi-sa* [want-NMLZ] and *tabe-sase-te yari-ta-sa* [eat-CAUS GER do.favor-want-NMLZ] are pronounced as one unit. The examples in (34) thus involve a syntax-morphology mismatch. The reality of the bracketed parts as syntactic constituents is ascertained by the case marking of the arguments in them. In (34a), the object 'money' is marked in the nominative because the stative predicate *hosi-i* 'want' by itself selects a nominative object. Likewise, in (34b) the direct and indirect objects ('food' and 'kids') are respectively marked in the accusative and the dative because the causative predicate *tabe-sase* 'make eat' takes these two objects. Since these clausal case markers normally do not show up with a nominal head – where the genitive case would be called for instead – it follows that the *-sa* nominalization takes place after the case markers on the arguments are

provided in syntax. Significantly, however, the suffix -*sa* cannot be separated from its host by a pause or any linguistic material.

In a nutshell, the nominalizer -*sa* appearing in (34)-type constructions has a syntactic status but is phonologically attached to its host elements. These two properties meet the characterization of clitics. Interestingly, -*sa* is not alone in allowing clausal case markers on arguments. Japanese in fact abounds in nominalizing suffixes that attach to syntactic phrases or clauses, a phenomenon which Seki (1977) referred to as *ku no hōsetsu* ('subsumption of a phrase in a word'). Typical examples look as follows (Seki 1977; Kageyama 1993).

(35) a. *Kanozyo wa* [*otoko ni   damas-are*]-*soo   da.*
        she   TOP [man   DAT deceive-PASS]-appearance COP
        'She looks like getting deceived by the man.'

   b. *Kare wa* [*nani   ka ii-ta*]-*ge   da.*
        he   TOP [something Q   say-want]-appearance COP
        'He looks as if he wants to say something.'

   c. *Titi   wa* [*yakusoku o   wasure*]-*gati   da.*
        father TOP [promise   ACC forget]-tendency COP
        'My father tends to forget his promises.'

Soo (da) 'look like' in (35a), conventionally categorized as an auxiliary verb in Japanese grammar, is historically a bound noun meaning 'appearance'. Likewise, -*ge* 'appearance' (35b) and -*gati* 'tendency' (35c) are bound nouns that function as copulative predicates in combination with the preceding verbs or adjectives. Like the suffix -*sa*, these bound nouns are phonologically attached to their host elements on the left and resist separation from them. Nevertheless, -*soo*, -*ge*, and -*gati* in (35) are syntactically independent of the preceding clauses because the arguments in the host clauses are marked with clausal case markers like dative and accusative. The clausal markers should be contrasted with the genitive marker *no* that regularly appears in noun phrase structure, as in (36).

(36) a. *opera no/*o   utai-te*
        opera GEN/*ACC sing-SUF
        'singer of operas'

   b. *sake   no/*o   nomi-ppuri*
        alcohol GEN/*ACC drink-way
        '(his) way of drinking alcohol'

The accusative marker is banned in (36) because *utai-te* 'singer' and *nomi-ppuri* 'way of drinking' are nouns derived in the lexicon, thus lacking verbal properties when they appear in syntactic structures.

Another peculiarity of the syntactically licensed suffixes is that they are not conditioned by the syntactic complexity of the clauses they attach to. Consider the examples in (37), where -*soo* and -*gati* take the whole coordinate clauses in brackets as their scope – a phenomenon called "suspended affixation", which is commonly attested in Turkish and other agglutinative languages (Kornfilt 2012).

(37) a. [*Kaze ga huki, yuki ga huri*]-*soo da.*
     [wind NOM blow, snow NOM fall]-appearance COP
     'It appears that wind will blow and snow will fall.'

   b. *Syatyoo wa* [*yakusoku o wasure, kaet-te simai*]-*gati da.*
     president TOP [promise ACC forget leave-GER finish]-tendency COP
     'The president tends to forget his promises and go home.'

If -*soo* and -*gati* were simple affixes, they should be attached to both of the coordinated verbs. This peculiarity, also observed with the nominalizer -*sa* discussed in (34), makes a striking contrast to ordinary lexical suffixes like -*te* '-er' and -*ppuri* 'way' in (36), which are totally incapable of accommodating coordinated verbs (Kageyama 1993).

(38) a. *[*utai, hiki*]-*te* ([sing (and) play]-er) 'person who sings and plays'

   b. *[*tabe, nomi*]-*ppuri* ([eat (and) drink]-way) '(his) way of eating and drinking'

The preceding observations lead us to postulate that -*sa*, -*soo*, -*ge*, -*gati*, and other similar suffixes are "syntactic suffixes" that select for a case-marked clause as their host, as schematically represented in (39).

(39)
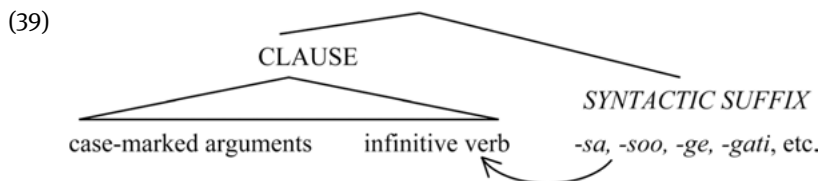
The question that immediately arises from this structure is how the suffix and its immediately preceding verb are fused into a morphological word. Does the verb raise to the suffix, or does the suffix lower to the verb? Kageyama (1982) opts for the lowering analysis on the grounds that the syntactic movement of raising the embedded verb to the suffix violates the Coordinate Structure Constraint (syntactic movement cannot apply only to part of a coordinate structure) when the complements are coordinated clauses. The lowering analysis is basically identical to "cliticization", as seen in the English genitive '*s*, which selects an NP syntactically

but is phonologically attached to its host, which immediately precedes it (cf. Zwicky and Pullum 1983).

(40) a. [$_{NP}$ John and his brother]'s car

  b. [$_{NP}$ the man who's standing there]'s wife

Consequently, what we have called "syntactic suffixes" are identified as clitics, or more precisely "enclitics".

The clitic analysis can be generalized to subordinate conjunctions and some verbal inflections that exhibit the "suspended affixation" effect. Observe the examples in (41) (cf. Kageyama 1982).

(41) a. [*sandoitti o tabe, koohii o nomi*]-**nagara**
      sandwich ACC eat coffee ACC drink while
      'while eating a sandwich and drinking coffee'

  b. [*haru ni nari, yuki ga toke*]-**reba**
      [spring DAT become snow NOM thaw]-if
      'if spring comes and snow thaws'

  c. [*hikooki ga epuron o hanare, hikoozyoo o ririku*]-**go**
      [airplane NOM apron ACC leave airport ACC take.off]-after
      'after the plane left the loading apron and took off the airport'

In (41), the subordinating morphemes *-nagara* 'while', *-reba* 'if', and *-go* 'after' syntactically take complement clauses but are phonologically unified with the immediately preceding words, so that the underlined parts are pronounced as a single phonological sequence. Because of this phonological behavior, these subordinators are identified as clitics. In (41a), the native subordinator *-nagara* 'while' takes a coordinate clause in its scope but attaches phonologically to the host verb *nomi* 'drink' given in the infinitive form. Other examples of this type are *-gatera* 'while', *-tutu* 'while', *-tuide ni* 'while', and *-zama ni* 'while at the same time'. While these morphemes are phonologically invariant no matter what verbs they attach to, the conditional morpheme *-reba* 'if' in (41b) alternates with *-eba*. Because of the allomorphy, it is regarded as an inflectional ending of the verb in traditional Japanese grammar, where *toke-reba* 'if (snow) thaws' is treated as a single word from the start. This view, shared by the lexicalist analysis today (e.g. Sells 1995), would encounter difficulty in explaining why the first verb *nari* 'become' in the first conjunct of (41b) is interpreted as a conditional ('if spring comes') even though it is not directly accompanied with *-reba*. On the other hand, clitics may exhibit allomorphic alternations as in the English genitive *'s*. The cliticization analysis would presumably be extendable to all the other verbal inflections (cf. Chapter 13 [Takezawa, this volume]

for the syntactic status of verbal inflections). The example in (41c) involves a Sino-Japanese subordinator morpheme *-go* 'after', which is appended to VNs. Also belonging to this category are *-tyuu* 'in the middle of', *-zen* 'before', and *-sidai* 'right after'.

Note that the characterization of syntactic suffixes as "clitics" does not apply to the nominalizing suffix *-kata* 'way', as in *piano no hiki-kata* [piano GEN play-way] 'how to play the piano'. Although it can attach to syntactically derived predicates like causative and passive, this suffix marks the arguments of the predicate it attaches to with the genitive instead of the accusative or other clausal markers. According to Kishimoto (2006), this is because *-kata* induces head-movement (or incorporation) of the base predicate, so that the whole vP phrase is turned into a noun phrase.

To summarize so far, Japanese offers sufficient motivation to differentiate lexical affixes from syntactic affixes. Although both are identified as bound affixes in resisting separation from their bases/hosts, they are distinct in syntactic behavior: Syntactic affixes can be attached to a clause while retaining clausal case marking on arguments, whereas lexical affixes cannot do so. Because of this, Hattori's notion of *fuzoku-keishiki* (bound affixes) should be split into "lexical" bound affixes and "syntactic" bound affixes, the former corresponding to archetypical affixes and the latter to "clitics". As far as we can see, all Japanese clitics are enclitics. Determiner-like prefixes of the W+ category such as *zen-* 'ex-, all' and *doo-* 'same', discussed in Section 2.3, are not clitics but are lexical prefixes (see also Hattori 1950 on this point). In the next subsection, we will move on to Hattori's *fuzoku-go* or non-independent words.

4.2 Particles and *na* ending as non-independent words

A topic of long-standing controversy in Japanese grammar is the morphological status of particles, which are generally classified into case particles (*ga* 'NOM', *o* 'ACC', *ni* 'DAT'), postpositions (*de* 'at, by', *kara* 'from', *yori* 'from, than', etc.), topic particle *wa*, focus particles (*mo* 'also, even', *sae* 'even', *made* 'even', *dake* 'only'), and sentence final particles (*yo* 'I tell you', *ka* 'Q', etc.). Are they words or bound morphemes? At first glance the answer seems simple. Since they are function words that cannot be used in isolation, they should be bound morphemes. In fact, Sells (1995) and other approaches in the strong lexicalist hypothesis (all word units are formed in the lexicon) treat the whole combination of a noun and a case particle as a word. However, there is no morphological evidence in favor of such a treatment.

Under the standard assumption, bound morphemes must cling to some other morphemes morphologically. This stipulation is contradicted by data that suggest that particles can stand alone. To save space, I focus here only on case particles, as these are the most likely candidates in the inventory of particles for analysis as bound morphemes. Below we argue, in the spirit of Hattori (1950) and in line with Vance (1993), that case particles are neither suffixes nor clitics, but are *fuzoku-go* (non-independent words).

First, Japanese case particles select a whole noun phrase rather than just a head noun. This is shown by the contrast between Japanese particles and Russian case declensions, pointed out by Vance (1993).

(42) a. *Taroo wa* [*sinbun   to   hon*] *o   yon-de   iru.*
Taro   TOP [newspaper and book] ACC read-GER is
'Taro is reading a newspaper and a book.'

    b. *Ivan chitayet   gazyet-u   i   knig-u.*
Ivan is-reading newspaper-ACC and book-ACC
'Ivan is reading a newspaper and a book.'

In the Japanese sentence (42a), the accusative particle *o* shows up only once, after the entire noun phrase consisting of two coordinated nouns 'a newspaper and a book', whereas in the Russian example (42b), the accusative ending *-u* is obligatorily attached to both 'newspaper' and 'book'. In Japanese, repetition of the particle *o* in (42a) results in an ungrammatical phrase like *⋆sinbun o to hon o* [newspaper ACC and book ACC]. Conversely in the Russian example, removal of the *-u* ending from either noun results in ungrammaticality, like *⋆gazyet i knig-u* [newspaper and book-ACC]. This indicates that the Russian *-u* is a bound affix but the Japanese *o* is not.

Can the Japanese particles be identified as clitics? Unlike the clitics discussed in Section 4.1 and unlike Russian declensional endings, Japanese case particles need not form a phonological unit with the preceding nouns. Concerning the pronunciation of case particles, it is well known that for some traditional speakers of the Tokyo dialect, the nominative particle *ga* has its initial velar [g] nasalized to [ŋ], and this fact is usually taken as a kind of word-internal sandhi that is indicative of the clitic status of *ga*. For speakers who have this pronunciation, however, the velar nasal is always used in this particle rather than alternating with [g] in particular environments. It is thus plausible to assume that for these speakers, the nominative particle is lexicalized as [ŋa]. In fact, Vance (1993) points out that there are instances of the velar nasal showing up even at word-initial position.

Vance (1993) presents a careful and comprehensive assessment of the morphological status of case and other particles in light of Zwicky and Pullum's (1983) and Zwicky's (1985) phonological, morphological, and syntactic criteria for clitics and concludes that Japanese particles are NOT clitics after all – a surprising conclusion for those who believe that they are. Below I reinforce the *fuzoku-go* (non-independent word) status of case particles by demonstrating that they can be separated from their host nouns.

First, case particles allow an additional material like *nado* 'etc.' and *izyoo* 'more than' before them, as in *hon nado o* [book etc. ACC] 'a book and other things'. Since *nado* and *izyoo* qualify themselves as non-independent words, case particles showing up after them are also non-independent words. Even more amazing is the fact observed by Hattori (1950: 15) that a parenthetical explanation can be put between a host noun and a particle, especially in written language, as shown in (43).

(43) *Kono syu no mondai – sore nituite wa notini kuwasiku ronzuru ga – o atukau.*
    this kind GEN problem – it about TOP later in.detail discuss but – ACC treat
    'I treat this kind of problem, about which I will discuss in detail later.'

Second, case particles can be optionally deleted by backward gapping in co-ordinate sentences.

(44) *Taroo wa yakyuu ~~ga~~ ~~suki de,~~ Ziroo wa sakkaa ‖ ga suki da.*
    Taro TOP baseball ~~NOM~~ ~~fond COP~~ Jiro TOP soccer NOM fond COP
    'Taro likes baseball, and Jiro soccer.'

In view of the principle of lexical integrity (the syntax cannot manipulate the internal structure of a word), the susceptibility of case particles to backward gapping, as in (44), indicates that they do not form a word with the preceding noun. Recall that gapped sentences are pronounced with a characteristic phonological break at the point marked "‖" in the second conjunct. Since the phonological break detaches the particle *ga* from the preceding noun *sakkaa* 'soccer' in (44), it is evident that the case particle is not cliticized to the noun on its left. Note incidentally that, even if the nominative particle *ga* is phonologically separated from the preceding noun, the velar nasal still occurs for those speakers who use it.

The qualification of case particles as non-independent words gains additional confirmation from the phonological behavior of mono-moraic nouns in the Kansai dialect. In this dialect, mono-moraic words such as *me* 'eye', *ha* 'tooth', and *ko* 'child' exhibit vowel lengthening when pronounced in isolation, as in [me:], [ha:], and [ko:]. The vowel lengthening does not take place inside compounds or derived words, as shown by *me-isya* 'eye doctor' (cf. *mee-isya) and *mae-ba* 'front tooth' (cf. *mae-baa). Put another way, the vowel lengthening in Kansai speakers takes place only at the right-hand edge of a phrase. Crucially, the vowels in such mono-moraic nouns are lengthened when followed by case particles, as in *mee ga* [eye NOM], *haa o* [tooth ACC], and *zibun no koo no* [self GEN child GEN] 'my child's'. This is a strong indication that there is a phrasal boundary between case particles and the nouns preceding them.

The foregoing discussion unequivocally shows that case particles are neither bound affixes (declensional suffixes) nor clitics. Is there, then, the possibility that they are full-fledged words? This possibility is easily denied by the fact that two or more case particles cannot be conjoined.

(45) a. *sensei [ga dewanaku o]
        teacher [NOM instead.of ACC]
        lit. 'the teacher (object) instead of the teacher (subject)'

    b. *sensei [kara toyuuyori ni]
        teacher [from rather.than to]
        'to rather than from the teacher'

The expressions in (46) are ungrammatical because the conjunctives *dewanaku* 'instead of' and *toyuuyori* 'rather than' do not apply to non-word-level morphemes. This demonstrates that *ga* and other case particles are not independent words.

Other instances of *fuzoku-go* that Hattori (1950) adduces include focus particles, sentence final particles, NP conjunctions like *ya* 'and', sentence conjunctions like *keredo* 'but' that occur after tensed clauses, the copula *da* and its allomorphs including the *na* ending occurring with ANs in adnominal position, *suru* used in combination with VNs, and many other morphemes. Here I focus on the status of *na*, leaving discussion on "VN *suru*" until Section 4.3.

As discussed in Chapter 2 (Kishimoto and Uehara, this volume) and Chapter 13 (Takezawa, this volume), the morphological status of the morpheme *na* used after ANs in prenominal position has been highly controversial in both descriptive and theoretical literature. One common view holds that it is an inflectional ending characteristic of ANs, on a par with the *-i* ending of native adjectives. This view is reflected in Masuoka and Takubo (1992), where native adjectives are referred to as *i*-adjectives and ANs as *na*-adjectives, following Mikami's (1953) idea that they constitute a larger group of adjectivals. Nishiyama (1999), by contrast, argues that both *na* and *-i* are variants of the copula. These two diametrically opposite views, however, both miss a vital difference in the morphological behavior of *-i* and *na*. Specifically, *na* can be separated from its host ANs in the same way that case particles can, whereas the affix *-i* cannot. Observe the sharp contrast between (46a) and (46b) involving the backward gapping of these two endings. As before, the symbol "||" indicates a phonological break.

(46) a. *Ken wa kooka* ~~*na tokei o kai,*~~ *Naomi wa anka ||*
       Ken TOP expensive *na* ~~watch~~ ~~ACC~~ ~~buy,~~ Naomi TOP cheap

       na tokei o kat-ta.
       na watch ACC buy-PST
       'Ken bought an expensive watch, and Naomi a cheap one.'

    b. **Ken wa taka-i* ~~*tokei o kai,*~~ *Naomi wa yasu ||*
       Ken TOP expensive-~~INFL~~ ~~watch~~ ~~ACC~~ ~~buy,~~ Naomi TOP cheap

       -i tokei o kat-ta.
       INFL watch ACC buy-PST
       'Ken bought an expensive watch, and Naomi a cheap one.'

In (46a), the bound morpheme *na* is pronounced as a separate constituent from the preceding AN *anka* 'cheap'. Such a separation is absolutely impossible with the bound affix *-i* in (46b).

The contrast between *na* and *-i* is further illuminated by the examples in (47) involving the conjunctive *toyuuyori* 'rather than, instead of'.

(47) a. [*mazime toyuuyori zittyoku*] *na*
      honest   rather.than sincere   ENDING
      'sincere rather than honest'

    b. *[*suzusi- toyuuyori samu*] *-i*
      cool   rather.than cold   ENDING
      'cold rather than cool'

In (47a), the *na* ending applies to the whole preceding AN phrase 'sincere rather than honest', in the same way that the case particle *ga* applies to conjoined NPs as in [*seito toyuuyori sensei*] *ga* ([student rather.than teacher] NOM) 'the teacher rather than the student'. These observations lead us to conclude that *na*, like the copula *da*, is a *fuzoku-go* (non-independent word). This conclusion will in turn count as evidence in support of Bloch (1946) and Okutsu (1978), who consider that the pre-nominal *na* ending of ANs is an allomorphic variant of the copula *da*, which itself is a non-independent word and has another allomorphic variant *ni* (identical in form with the dative particle *ni*) in adverbial usage, as in *kenkoo ni naru* [healthy COP become] 'become healthy', and *de* in conjunctive usage, as in *kenkoo de* [healthy COP] 'healthy and'. The total ungrammaticality of (47b), on the other hand, confirms that the *-i* ending on native adjectives is a bound affix, and not a copula.

    By way of further illustrating the notion of *fuzoku-go*, I will briefly touch on the morphological status of *hatu* 初 'the first time', which has two semantically similar but functionally distinct uses as a prefix attached to the left of a host noun (48a) and as a *fuzoku-go* appearing after a host noun (48b).

(48) a. *hatu-koi* [first-love] 'one's first love'

    b. *sekai hatu* [world first] 'for the first time in the world'

These two are phonologically different in that the prefix *hatu-* in (48a) makes a single accentual unit with *koi* 'love', as in *haTU-KOI*, while the morpheme *hatu* in (48b) carries its own accent, as in *SEkai HAtu*. It is the latter "suffix-like" *hatu* that qualifies as *fuzoku-go*. Observe that it is subject to backward gaping in (49).

(49) *Kore wa sekai* ~~*hatu*~~ ~~*no*~~ ~~*hatumei*~~ ~~*de,*~~ *are wa Nihon* ||
    this TOP world ~~first~~ ~~GEN~~ ~~invention~~ ~~COP~~ that TOP Japan

    hatu no hatumei da.
    first GEN invention COP
    'This is the world's first invention, and that one Japan's first invention.'

The status of *hatu* as a non-independent word is confirmed by the 'rather than' test, as in [*Nihon toyuuyori sekai*] *hatu* ([Japan rather.than world] first.time) 'the first time in the world rather than in Japan'.

In this section, we have zoomed in on two cases where morphological entities participate in the formation of syntactic structure: clitics and non-independent words. Table 3 summarizes the characteristics of these two kinds of morphological object as compared with lexical suffixes and freestanding words.

Table 3: Characteristics of suffixes, clitics, and non-independent words

	free words	*fuzoku-go*	clitics	bound affixes
syntactic status	Yes	Yes	Yes	No
morphological independence	Yes	Yes	No	No
freestanding	Yes	No	No	No

Both clitics and non-independent words (*fuzoku-go*) have their own position in syntactic structure as is clear from the fact that they take phrasal or clausal categories, participating in the so-called suspended affixation phenomena. Clitics, however, are attached to their hosts in phonological structure, whereas non-independent words remain *in situ* without being adjoined to their hosts. Since they are generated outside phrases or clauses and hence do not manipulate the internal structure of a given word, it is concluded that they are irrelevant to the lexical integrity principle stated in (1). Finally, the significance of Hattori's (1950) notion *fuzoku-go* should be particularly underscored as it is a category absent from European linguistics. The concrete examples of *fuzoku-go* or non-independent words all show up at the right-hand periphery of phrases and clauses, and this property is considered to fall out from the characteristic of Japanese as a left-branching agglutinative language.

4.3 On the wordhood of "VN *suru*"

Discussion of lexical integrity in Japanese would not be complete without taking up the long-standing issue of the morphological status of "VN *suru* (do)" sequences in the so-called light verb construction. While the syntactic and semantic properties of this construction are discussed in depth in Chapter 12 (Miyamoto and Kishimoto, this volume), the morphological status of VN *suru* complexes like *seityoo suru* [growth do] 'to grow' and *bunseki suru* [analysis do] 'to analyze' has long been an unsettled question in both descriptive Japanese grammar and theoretical morphology. On one view, advocated by Matsumoto (1996), Iida and Sells (2008), and others, "VN *suru*" is a mere juxtaposition of two words. Booij's (2010) Quasi-incorporation analysis also holds that N^0 (i.e. VN) is adjoined to V^0 (i.e. *suru*) to form a V^0 constituent of the form [$_{V0}$ VN *suru*]. The claim that each of the VN and *suru* has morphosyntactic

independence is easily shown by examples like (50a), where focus particles like *sae* 'even' and *mo* 'also' can intervene between the two elements.

(50) a. *Sono toki kyodai na tunami ga hassei (sae/mo) si-ta.*
        that time massive COP tsunami NOM occur (even/also) do-PST
        'At that time, it even/also happened that a massive tsunami occurred.'

   b. *tunami no hassei no* [*si-kata*]
        tsunami GEN occur GEN [do-way]
        'the way the tsunami occurred'

   c. **tunami no* [*hassei si-kata*]
        tsunami GEN [occur do-way]

In (50a), an unaccusative VN *hassei* 'occur' is intentionally employed to preclude the possibility that the sequence *hassei sae/mo si-ta* [occur even/also do-PST] results from a simple ellipsis of the accusative particle *o* before the focus particles.[2] The two-word analysis gains further support from the ungrammaticality of (50c), where the whole of VN *suru* complex fails to undergo nominalization by the suffix *-kata*; instead, only the verb *suru* is subject to nominalization as in (51b) (Kageyama 1993: 206). The contrast between (51b) and (51c) shows that the sequence of VN and *suru* does not make up a single verb at least at the stage where the syntactic *-kata* nominalization takes place.

    The possibility of particle insertion between VNs and *suru* as in (50a) also led Hattori (1950) to conclude that *suru* belongs to the class of *fuzoku-go*. However, evidence is also available that argues against the status of *suru* as a non-independent word. Compare (51a) with (51b) involving backward gapping (cf. Kageyama 1993: 261).

(51) a. *A-san wa tuma to bekkyo sae ~~si~~, B-san wa rikon sae* || *si-ta.*
        A-Mr TOP wife with separate even ~~do~~, B-Mr TOP divorce even do-PST
        'Mr A even got divorced, and Mr B even got separated from his wife.'

   b. **A-san wa tuma to bekkyo ~~si~~, B-san wa rikon* || *si-ta.*
        A-Mr TOP wife with live.separately ~~do~~, B-Mr TOP divorce do-PST
        'Mr A got divorced, and Mr B got separated from his wife.'

In (51a), the emphatic particle *sae* 'even' intervenes between VNs and *su(ru)*. As before, the symbol "||" in the second conjunct shows that there is a short phonological break between *sae* and *si-ta*. The phonological break demonstrates that

---

2 In Japanese, the accusative particle *o* on any noun generally drops out when followed by a focus particle, and the "VN *suru*" in light verb constructions has an alternative realization with the accusative particle *o* attached to the VN, as in "VN *o suru*". In accordance with Burzio's generalization (Only verbs that have external argument may assign accusative case), the VN can bear the accusative marker if it is transitive or unergative (i.e. has an external argument), but not if it is unaccusative. See Chapter 12 (Miyamoto and Kishimoto, this volume) for references on this issue.

si-ta 'did' is a separate constituent from *rikon sae* [divorce even]. Since the sequence *VN sae suru* [VN even do] is not a word, the grammaticality of (51a) is well expected. On the other hand, the sentence in (51b), designed to have the same construction, but this time without an intervening focus particle, is judged ill-formed. It is utterly unnatural to pronounce the sentence with the characteristic phonological break at ||, thus separating *rikon* 'divorce' from *si-ta* 'did' in the second conjunct. Note that the same sentence would be acceptable on an entirely different interpretation where no phonological break is put at || and the VN *bekkyo* 'live separately' in the first conjunct closes off the clause, resulting in a sentence fragment without tense.

The significance of the failure of gapping in the VN *suru* construction as in (51b) carries weight if we compare it with the successful application of the same deletion to a superficially analogous sequence of a mimetic adverb followed by *suru*. As shown by Kageyama (2007), it is possible in (52) to delete only the verb *si* in the first conjunct, leaving its partner word (*gakkari* 'disappointed') behind.

(52) *Sore o kii-te, ani wa **gakkari** ~~si,~~ otooto wa*
     that ACC hear-GER big.brother TOP MIM do little.brother TOP

     ***ninmari** || si-ta.*
     MIM do-PST
     'Hearing that, the big brother was disappointed, and the little brother delighted.'

This shows that the complex of a mimetic word and *suru* does not make up a word. The contrast between (51b) and (52) strongly suggests that the sequence of VN and *suru* makes up a kind of word at the stage where the phonological rule of backward gapping takes place, although it is composed of two separate words at the level of syntax. In other words, *suru* does not qualify as a *fuzoku-go* because it cannot be separated from the preceding VN by a pause, according to Hattori's (1950) criterion.

Provided that a VN and *suru* are separate words at some level of syntax but are combined into one word at a later stage, we must ask exactly what kind of word formation process is responsible for this compounding. Perhaps the least likely possibility is that *suru* is cliticized to its host VN. Observe the ungrammaticality of (53), where *suru* is attached to phrasal categories (cf. Kageyama 1993: 262).

(53) a. **Watasi wa sensei ni [aisatu to onegai]-si-ta.*
       I TOP teacher DAT [greeting and request]-do-PST
       'I greeted my teacher and asked him a favor.'

     b. **Watasi wa [Ainu-go no benkyoo]-si-ta.*
       I TOP [Ainu-language GEN research]-do-PST
       'I studied the Ainu language.'

(53a) involves a phrase coordinating two VNs, and (53b) a phrase with a genitive modifier. If *suru* is a clitic, it should be able to freely attach to phrasal categories as we saw earlier with case particles in (42a). Notice that (53a) and (53b) become

perfectly grammatical if *sae* 'even' or the accusative particle *o* is added after the VN phrases to make *suru* a separate word.

Another possible analysis, first suggested by Kageyama (1977), is that a VN is "incorporated" with *suru* in syntax. However, if VN incorporation takes place in the syntactic cycle, it is obviously at odds with the facts regarding *-kata* nominalization observed in (50c). A slightly different analysis is suggested by Kageyama (1993: 360), according to which the compounding or incorporation takes place at "post-syntactic" structure, namely at a level in syntax after lexical accent is assigned (cf. Shibatani and Kageyama 1988; Chapter 7 [Kageyama, this volume]). Kageyama's (1993) suggestion is based on the accentuation of "VN *suru*".

(54)	**VN alone**	**VN-*suru***	**VN *o suru***
a.	*KYO-ka* 'permit'	*KYO-ka-su-ru*	*KYO-ka o su-ru*
	H-L	H-L-L-L	H-L L L-L
b.	*zya-MA* 'disturb'	*zya-MA-SU-RU*	*zya-MA O SU-RU*
	L-H	L-H-H-H	L-H H H-H
c.	*i-KI-ki* 'come and go' *i-KI-ki-su-ru*		*i-KI-ki o su-ru*
	L-H-L	L-H-L-L-L	L-H-L L L-L

As shown in (54), the accent pattern of a whole VN-*suru* sequence (the middle column) is determined by that of the VN element (the left column), and this pattern is parallel to the accentuation of accusative-marked phrases (the right column). This correlation indicates that VN-*suru* takes over the accent pattern of its phrasal counterpart. Since manifestation of phrasal rather than lexical accent is a hallmark of post-syntactic compounds, the formation of VN-*suru* words can also be considered to take place in post-syntactic structure.

The post-syntactic derivation can not only accommodate all the relevant phenomena discussed above but will also make the correct prediction about new constructions like the following (Kageyama 1993: 345).

(55) a. *undoo | si-tari   si-nakat-tari*
         exercise   do-CONJCT do-NEG CONJCT
         '(I) sometimes get exercise and sometimes don't'

     b. *tunami   ga   hassei | si-te-mo   si-naku-te-mo*
         tsunami NOM occur   do-GER-also do-NEG GER-also
         'no matter whether a tsunami occurs or not'

In (55), VNs *undoo* 'exercise' and *hassei* 'occur' are followed by a complex of two *suru* verbs, one in the affirmative and the other in the negative, and a slight pause can be put after the VNs at the point marked " | ". This and other facts leads Kageyama (1993: 345) to analyze the constructions in (55) along the lines of the bracketed structure in (56).

(56) *undoo*　　[ᵥ₊　*si-tari*　　　　*si-nakat-tari*]
　　　exercise　[　　do-CONJCT　do-NEG CONJCT]

In this structure, *si-tari si-nakat-tari* makes up a Word plus unit, to which the VN *undo* is attached. The special W⁺ category here, presumably available only for the idiomatic frame involving pairs of positive and negative verbs or of antonymous verbs such as *it-tari ki-tari* 'go back and forth', could be extended to complex verbs involving syntactic suffixes like passive and causative.

(57)　a.　*hihan*　　[ᵥ₊　*s-are-tari*　　　　*s-are-nakat-tari*]
　　　　　criticize　[　　do-PASS CONJCT　do-PASS NEG CONJCT]
　　　　　'be sometimes criticized and sometimes not'

　　　b.　*undoo*　　[ᵥ₊　*s-ase-tari*　　　　*s-ase-nakat-tari*]
　　　　　exercise　[　　do-CAUS CONJCT　do-CAUS NEG CONJCT]
　　　　　'sometimes make (a person) get exercise and sometimes don't'

The complex amalgamations of VN-*suru* and syntactic suffixes in (57) demonstrate that the VN incorporation takes place at post-syntactic structure after passive and causative verbs are formed in the syntax.

5 Conclusion and future research perspectives

This chapter discussed the interaction of morphology and syntax with particular attention to the principle of lexical integrity, and found that the applicability of this principle is not uniform but varies according to the nature and range of individual phenomena. It is not appropriate to conclude that the principle is correct or incorrect holistically; instead, it is necessary to disintegrate the principle as stated in (1) into at least two components, as proposed independently by Booij (2008) and Kageyama (2009).

(58)　Disintegration of lexical integrity
　　　a.　No syntactic access: The syntax cannot have access to only a part of word structure.

　　　b.　No syntactic deformation: Syntactic rules cannot affect the internal structure of words

　　　The "no syntactic access" constraint in (58a), corresponding to the first half of Anderson's (1992) principle in (1), is referred to as "non-accessibility" by Booij (2008) and as "syntactic unanalyzability" by Kageyama (2009), whereas the "no syntactic deformation" constraint in (58b), corresponding to the second half of Anderson's principle and comprising the No Phrase Constraint, is referred to as "non-interruptability" by Booij (2008) and "syntactic indeformability" by Kageyama (2009). Both Booij and Kageyama conclude on independent grounds that "no syntactic deformation" is a universal constraint while "no syntactic access" is not.

In Section 2, we discussed three Japanese phenomena that were largely concerned with the "no syntactic deformation" constraint, namely, genitive compounds, phrasal compounds, and Word plus, and concluded that genitive compounds and Word plus are manageable in the architecture of morphology with slight modification of the existing system; phrasal compounds, however, are in conflict with the constraint although their exceptional behavior appears to be controllable by the lexical properties of head and non-head constituents. It should be stressed that the categories of syntax and those of morphology are not completely disparate but are continuous, with the W⁺ category sitting at their intersection. The W⁺ category is not peculiar to Japanese but appears to be shared by other languages, as suggested by Kageyama (2001, 2013).

Section 3 brought forth phenomena relevant to the "no syntactic access" constraint and found that the W⁺ prefix *doo-* 'aforementioned' is a true counterexample to the Anaphoric Island Constraint. Sentences involving this prefix can freely refer into and out of complex word structure, thanks to its intrinsic referential function. Finally, Section 4 tackled clitics and non-independent words, which show up at the right-hand periphery of phrasal and clausal categories. Since they neither access nor deform the internal structure of existing words, they are simply outside the scope of the two constraints in (58). All in all, the "no syntactic access" constraint is not universally supported, whereas the "no syntactic deformation" constraint appears universally valid, with the caveat that phrasal compounds are manageable exceptions.

To see the validity of "no syntactic deformation", we might consider two cases of syntactic movement that pertain to morphology. One is movement of a material into a word to create a complex word, and the other is movement of a material out of an existing complex word. Clearly, these two cases have different implications for lexical integrity. The former brings about no effect on the "no syntactic deformation" constraint because its function is to create rather than destroy a word structure. In particular, syntactic movement of a head to another head to create a complex predicate is recognized as Incorporation (Baker 1988), and we indeed saw in Section 4.3 that VN is "incorporated" with *suru* in the light verb constructions (see Chapter 7 [Kageyama, this volume] for other examples). On the other hand, the latter case, where a part of a word is syntactically removed from the rest of the word – often called "excorporation" – remains to be explored. Is it ever possible for a morpheme to float away from the host word in syntactic structure? Kitagawa (1986) entertains the idea of excorporation in Logical Form (LF) in analyzing the peculiar meaning of the native prefixes *ko-* 'small' and *oo-* 'big' that appear in certain idioms like *ko-waki ni kakaeru* [small-arm DAT hold] 'carry lightly under one's arm'. Specifically, he argues that these prefixes, which are morphologically attached to nouns (e.g. *ko-waki* [small-arm]) but are semantically interpreted as degree modifiers taking scope over the verb phrases (e.g. 'hold lightly'), move away from their base nouns to a higher position of VP in LF. Since, however, the prefixes in question are not productive, these special constructions could be plausibly disposed of as idioms without having recourse to excorporation.

Another likely candidate for excorporation might be the Sino-Japanese prefix *tyoo-* [tʃoː] 'super', which is highly productive. Over the recent years, the prefix *tyoo* 'super' has been used or overused by younger speakers as an emphatic adverb in casual speech. While this prefix normally selects noun bases, it is also used with predicates such as adjectives, ANs, and verbs, as shown in (59). Note that this instance of *tyoo* is pronounced with its own accent and is followed by a short break – a trait of the W⁺ prefixes discussed in Section 2.3.

(59) a. *tyoo kakkoii* 'super cool',   *tyoo dasai* 'super uncool'

   b. *Tyoo tukare-ta*. 'I'm super tired.'   *Tyoo arui-ta*. 'I walked a lot.'

Examples like those in (59), where *tyoo* is attached to adjectives and verbs, are not too much of a problem, because they are simple extension of base categories from nouns to verbs and adjectives. The same prefix, however, can even be attached to predicate phrases such as VP and AP, as exemplified by (60).

(60) a. *Tyoo* [$_{VP}$ *hara   (ga)   het-ta*].
       super [stomach (NOM) get.smaller-PST]
       'I'm dead hungry.'

   b. *Tyoo* [$_{AP}$ *kimoti (ga)   waru-i*].
       super [feeling (NOM) bad-PRS]
       'I'm very unpleasant.'

   c. *Tyoo* [$_{VP}$ *ki   ni   nar-u*].
       super [mind DAT become-PRS]
       'I'm very worried (about it).'

   d. *Tyoo* [$_{VP}$ *sake (o)   non-da*].
       super [alcohol (ACC) drink-PST]
       'I drank a lot.'

In light of examples like (59) and (60) we might venture the hypothesis that the prefix *tyoo-*, which was originally attached to a bound morpheme (*tyoo-zin* 'superman') and a word (*tyoo-tokkyuu* 'super-express'), was extended to the W⁺ category (*tyoo | oomono* 'super big shot') and has acquired the status of *fuzoku-go* or non-independent word in (60), although this last usage is accepted only in younger speakers' casual speech. Even in this novel usage, however, the base of prefixation is maximally limited to predicate phrases (VP, AP) and can never extend to a whole sentence (*v*P or IP), as shown in (61).

(61) **Tyoo kinoo   watasi (wa) kimoti yokat-ta*.
    super yesterday I   TOP feeling good-PST
    'I felt super yesterday.'

To sum up, the prefix *tyoo* appears to have acquired the versatility of attaching not only to Root, Word, and Word plus, but also to predicate phrases by expanding its scope from a morphological unit to a syntactic phrase. The growth of *tyoo* from a

bound affix to a non-independent word appears incompatible with the unidirectionality hypothesis of grammaticalization, according to which a word may change to a bound affix but not vice versa. As Hopper and Traugott (2003) caution, however, change in the opposite direction, called "degrammaticalization", may happen sporadically. The *fuzoku-go* usage of *tyoo* in (60) would be regarded as a case of degrammaticalization. Since this usage of *tyoo* is a result of historical change, however, it does not affect the validity of the "no syntactic deformation" constraint of lexical integrity, because the prefix is not physically "removed" from its host by any kind of syntactic movement rule.

Acknowledgments

References

Anderson, Stephen. 1992. *A morphous morphology*. Cambridge: Cambridge University Press.

Aoyagi, Seizō. 1969. A demarcative pitch of some prefix-stem sequences. *Study of Sounds* 12. 241–247.

Baker, Mark. 1988. *Incorporation: A theory of grammatical relation changing*. Chicago: University of Chicago Press.

Bloch, Bernard. 1946. Studies in colloquial Japanese I: Inflection. *Journal of the American Oriental Society* 66. 97–109. Reprinted in Roy A. Miller (ed.). 1969. *Bernard Bloch on Japanese*. New Haven: Yale University Press.

Booij, Geert. 2008. Lexical integrity as a formal universal: A constructionist view. In Sergio Scalise, Elisabetta Magni and Antonietta Bisetto (eds.), *Universals of language today*, 83–100. Dordrecht: Springer.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Botha, Rudolf P. 1983. *Morphological mechanisms*. Oxford: Pergamon Press.

Bresnan, Joan and Sam Mchombo. 1995. The lexical integrity principle. *Natural Language & Linguistic Theory* 13. 181–254.

Di Sciullo, Anna Maria. 2005. *Asymmetry in morphology*. Cambridge, MA: MIT Press.

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Hattori, Shirō. 1950. Fuzoku-go to fuzoku-keishiki [Subordinate words and subordinate forms]. *Gengo Kenkyu* 15. 1–26. Reprinted in Shirō Hattori. 1960. *Gengogaku no hōhō* [Methods of linguistics], Tokyo: Iwanami Shoten.

Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.

Iida, Masayo and Peter Sells. 2008. Mismatches between morphology and syntax in Japanese complex predicates. *Lingua* 118. 947–968.

Kageyama, Taro. 1977. Incorporation and Sino-Japanese verbs. *Papers in Japanese Linguistics* 5. 117–155.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–58.

Kageyama, Taro. 1989. The place of morphology in the grammar: Verb-Verb compounds in Japanese. In Geert Booij and Jaap van Marle (eds.), *Yearbook of morphology* 2. 73–94. Dordrecht: Foris.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1999. *Keitairon to imi* [Morphology and semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2001. Word plus: The intersection of words and phrases. In Jeroen van de Weijer and Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*, 245–276. Berlin: Mouton de Gruyter.

Kageyama, Taro. 2007. Explorations in the conceptual semantics of mimetic verbs. In Bjarke Frellesvig, Masayoshi Shibatani and John Smith (eds.), *Current issues in the history and structure of Japanese*, 27–82. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber and Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kageyama, Taro. 2013. Post-syntactic compounds and semantic head-marking in Japanese. In Bjarke Frellesvig and Peter Sells (eds.), *Japanese/Korean Linguistics* 20, 363–382. Stanford: CSLI Publications.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun-incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Lexical strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kishimoto, Hideki. 2006. Japanese syntactic nominalization and VP-internal syntax. *Lingua* 116. 771–810.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kitagawa, Yoshihisa. 1986. More on bracketing paradoxes. *Linguistic Inquiry* 17. 177–183.

Kornfilt, Jaklin. 2012. Revisiting "suspended affixation" and other coordinate mysteries. In Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro and Cecilia Poletto (eds.), *Functional heads: The cartography of syntactic structures, volume 7*, 181–196. Oxford: Oxford University Press.

Lapointe, Steven. 1980. *A theory of grammatical agreement*. Amherst, MA: University of Massachusetts dissertation.

Lieber, Rochelle. 1992. *Deconstructing morphology*. Chicago: University of Chicago Press.

Lieber, Rochelle and Sergio Scalise. 2006. The lexical integrity hypothesis in a new theoretical universe. *Lingue e linguaggio* 1. 7–32.

Martin, Samuel E. 1975. *A reference grammar of Japanese*. New Haven: Yale University Press.

Masuoka, Takashi and Noriyuki Takubo. 1992. *Kiso nihongo bunpō*, revised ed. Tokyo: Kurosio Publishers.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford: CSLI Publications and Tokyo: Kurosio Publishers.

Mikami, Akira. 1953. *Gendai gohō josetsu* [Prolegomena to modern usage]. Tokyo: Tōkō Shoin.

Miller, D. Gary. 1993. *Complex verb formation*. Amsterdam: John Benjamins.

Miyamoto, Tadao and Hideki Kishimoto. this volume. Chapter 12: Light verb constructions with verbal nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Namiki, Takayasu and Taro Kageyama. this volume. Chapter 6: Word structure and headedness. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer.

Nishiyama, Kunio. 1999. Adjectives and the copulas in Japanese. *Journal of East Asian Linguistics* 8. 183–222.

Nishiyama, Yuji. this volume. Chapter 18: Complement-taking nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Okutsu, Keiichirō. 1978. *"Boku wa unagi da" no bunpō* [The grammar of "I am an eel"]. Tokyo: Kurosio Publishers.

Poser, William. 1990. Word-internal phrase boundary in Japanese. In Sharon Inkelas and Draga Zec (eds.), *The phonology syntax connection*, 279–288. Chicago: University of Chicago Press.

Postal, Paul. 1969. Anaphoric islands. In Robert I. Binnick, Alice Davison, Georgia M. Green and Jerry L. Morgan (eds.), *Papers from the Fifth Regional Meeting Chicago Linguistic Society*, 205–239.

Roeper, Thomas. 1987. Implicit arguments and the head-complement relation. *Linguistic Inquiry* 18. 267–310.

Roeper, Thomas and Muffy E.A. Siegel. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9. 199–260.

Seki, Kazuo. 1977. *Kokugo fukugōdōshi no kenkyū* [A study of Japanese compound verbs]. Tokyo: Kasama Shoin.

Selkirk, Elizabeth. 1982. *The syntax of words*. Cambridge, MA: MIT Press.

Sells, Peter. 1995. Korean and Japanese morphology from a lexicalist perspective. *Linguistic Inquiry* 26(2). 277–325.

Shibatani, Masayoshi and Taro Kageyama. 1988. Word formation in a modular theory of grammar. *Language* 64. 451–484.

Sproat, Richard. 1988. On anaphoric islandhood. In Michael Hammond and Michael Noonan (eds.), *Theoretical morphology*, 291–301. San Diego: Academic Press.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Sugioka, Yoko. 1989. Haseigo niokeru dōshi-sosei no uketsugi [Inheritance of verb features in derived nouns]. In Susumu Kuno and Masayoshi Shibatani (eds.), *Nihongogaku no shin tenkai* [New developments of Japanese linguistics], 167–185. Tokyo: Kurosio Publishers.

Sugioka, Yoko and Takane Ito. this volume. Chapter 10: Derivation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Takezawa, Koichi. this volume. Chapter 13: Inflection. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Taylor, John. 1996. *Possessives in English*. Oxford: Clarendon Press.

Vance, Timothy J. 1993. Are Japanese particles clitics? *Journal of the Association of Teachers of Japanese* 27(3). 3–33.

Ward, Gregory, Richard Sproat and Gail McKoon. 1991. A pragmatic analysis of so-called anaphoric islands. *Language* 67. 439–472.

Yamanashi, Masa-aki. 1992. *Shōō to suiron* [Anaphora and inference]. Tokyo: Kurosio Publishers.

Yumoto, Yoko. 2005. *Fukugōdōshi haseidōshi no imi to tōgo* [Semantics and syntax of compound verbs and complex verbs]. Tokyo: Hituzi Syobo.

Zwicky, Arnold. 1985. Clitics and particles. *Language* 61. 283–305.

Zwicky, Arnold and Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59. 502–513.

Wesley M. Jacobsen

15 Lexical meaning and temporal aspect

1 Introduction

Natural language imposes structure on the world in which we live by packaging space and time into individual units called events. Central to this process is the role played by the predicate, the basic linguistic unit from which sentences are built. Predicates structure events in two fundamental ways. One is by selecting a set of entities that are understood to participate in an event, together forming what is called the argument structure of a predicate. In sentence (1), for example, the predicate *ireru* 'put in' requires three entities that must be present for the event it expresses to make sense, represented by the nouns *Suzuki-sensyu* 'athlete Suzuki', *booru* 'ball', and *gooru* 'goal'.

(1) *Suzuki-sensyu   ga   booru   o   gooru   ni   ire-ta.*
    Suzuki-athlete   NOM   ball   ACC   goal   LOC   put.in-PST
    'Suzuki scored a goal (lit. put the ball into the goal).'

Another is by assigning to the event a temporal structure that describes how it unfolds in time. The temporal structure assigned to the event in (1) by *ireru* and its accompanying nouns consists of two parts, an activity of the athlete Suzuki that has some duration in time, however short, and a terminal point of the activity corresponding to the moment in time the ball enters the goal.

This chapter is concerned with how events are expressed in Japanese as unfolding in time in this way, focusing on the role played by individual predicates in expressing such meaning, that which defines the subject matter of the linguistic study of (temporal) aspect. The question of how events are seen to unfold in time is closely bound up with the question of how events are situated in time, that which forms the subject matter of the linguistic study of tense. Tense is the study of how languages order events in time *relative* to other events, most importantly the time of speech, but also to other events that may be present in the linguistic context (Ogihara 1999). Languages may do this in various ways, and can differ in whether such ordering must be overtly marked on the predicate or not. Japanese is a language that allows no option in this regard. Every predicate must be marked by one of a two-way opposition of tense forms: *-ta* vs. *-ru* (or its allophonic variant *–u*) for verbs, *-katta* vs. *-i* for adjectives, and *-ta* vs. *-Ø* for the copula. In main clauses, these order an event with respect to the time of speech, either earlier than it or not earlier than (at the same time as or later than) it, as illustrated in (2).

(2) a. *Sanzi   ni   Kyooto   ni   tui-ta/tuk-u.*
        3:00   at   Kyoto   LOC   arrive-PST/arrive-NONPST[1]
        'We arrived/will arrive in Kyoto at 3:00.'

     b. *Yasumi   no   hi   wa   uti   ni   i-ta/i-ru.*
        vacation   GEN   day   TOP   home   LOC   exist-PST/exist-NONPST
        'I was/am at home on vacation days.'

Aspect is concerned with the internal temporal structure of the event itself, apart from how it is ordered with respect to other events, including such features as the presence or absence of a transition to a new state in the course of the event. The verb *tuku* 'arrive' in (2a), for example, denotes an event whereby a transition occurs in the location of the subject from not being in Kyoto to being in Kyoto, while the verb *iru* 'be, exist' in (2b) denotes an event (situation) that is unchanging throughout the relevant time frame – the same situation obtains no matter what subpart of the time frame is considered. Anything that has internal structure, however, is made up of subparts that are ordered with respect to each other, which means that considerations of ordering (tense) cannot be neatly distinguished from considerations of structure (aspect) (Jacobsen 2005).

Japanese exhibits numerous auxiliary-like forms that express a variety of aspectual meanings, the variety being even greater in historically earlier periods of the language (Classical Japanese) and in Western dialects of contemporary Japanese. For this reason, there has been a long-standing interest in these forms among native Japanese grammarians, who have generated an extensive corpus of research on the topic, some of which is discussed in Chapter 13 (Takezawa, this volume), dealing with inflectional endings, and Chapter 11 (Nakatani, this volume), dealing with concatenated V-*te* V complex predicates, as well as in companion volumes to this volume, such as the *Handbook of Japanese semantics and pragmatics* and *Handbook of Japanese dialects*. The present chapter, making up as it does a part of *Handbook of Japanese lexicon and word formation*, is concerned with the basic properties of "lexical aspect" – aspect that is inherent in the meaning of verbs and other predicates, and takes as a point of departure a critical examination of the work of Haruhiko Kindaichi (1950), widely acclaimed as a seminal work on lexical aspect in Japanese. This chapter will attempt to clarify the role played by lexical aspect not only as an organizing principle in the lexicon, but also in constraining the alternation and realization of noun arguments of predicates in syntactic constructions.

---

1 Note that the inflectional suffixes *ta* and *ru* are rendered in the literal glosses of this chapter as PST (past) and NONPST (non past), respectively, in contrast to the convention followed in the other chapters of this volume of rendering the *ru* form as PRS (present). This is intended to reflect the possibility of either a literal present or future interpretation in this form.

2 Grammatical (viewpoint) vs. lexical (situation) aspect

How exactly is temporal meaning "structured" in languages and where do we look to find such meaning? Broadly speaking, such meaning can be found in two places. One is where the meaning is explicitly expressed in a grammatical form, such as an affix, auxiliary verb, or compound verb attached to the predicate, or even an adverb or other independent element in the clause to which the predicate belongs. This "grammatical aspect" corresponds to what Smith (1997) has called "viewpoint aspect", as these forms impose additional aspectual meaning – a particular "viewpoint" – on a temporal structure that already exists. In (3), for example, the verbs *hazimeru* 'begin', *tuzukeru* 'continue', and *owaru* 'end' appear as the second element in a compound V1-V2 structure to highlight the beginning, intermediate, and final stages of the event of *beru ga naru* 'the bell rings'.

(3) *Syuugyoo   no   beru   ga   nari-{hazime-ta/tuzuke-ta/owat-ta}.*
     end.of.class   GEN   bell   NOM   ring-{begin-PST/continue-PST/finish-PST}
     'The bell signaling the end of class began/continued/finished ringing.'

V2 as the second element of a verb concatenation structure is the most common means employed in Japanese for expressing viewpoint aspect, whether of the compound type V1-V2 seen in (3) or V1-*te* V2 gerund structure in (4), where V2 in the V1-*te iru* pattern imposes progressive aspect on the event (see Nakatani 2004; Chapter 11 [Nakatani, this volume]).

(4) *Kodomo   ga   niwa   de   ason-de   iru.*   (*asob + te + iru*)
     children   NOM   yard   LOC   play-GER   be.NONPST
     'The children are playing in the yard.' (progressive)

But viewpoint aspect can also be expressed by grammaticalized patterns other than verb concatenation, such as the V-*ta koto ga aru* pattern in (5), which imposes "experiential aspect" – a variety of perfect aspect – on the event expressed by the associated verb.

(5) *Huzi-san   ni   nobot-ta   koto   ga   ar-u.*
     Fuji-Mt.   GOAL   climb-PST   COMP   NOM   exist-NONPST
     'I have (the experience of having) climbed Mt. Fuji.'

It can also be imposed by elements independent of the predicate itself, such as adverbs co-occurring with a predicate, as seen in the distinction between a simple future and habitual reading triggered by the different adverbs in (6).

(6) *Kotosi   no   natu/mai-tosi   natu   Nihon   e   kaer-u.*
    this.year GEN summer/every-year summer Japan GOAL return-NONPST
    'I'm returning to Japan this summer/I return to Japan every summer.'

In those cases where aspectual meaning is encoded by the V2 form in a verb concatenation pattern, the original meaning of the verb is typically transparent to some degree, making identification of the aspectual meaning fairly straightforward, although various degrees of "semantic bleaching" may occur, where features of the meaning or behavior of the original verb are lost. In (3), for example, the forms *hazimeru* 'begin' and *tuzukeru* 'continue' are originally transitive, requiring that an object argument slot be filled, but that feature of these verbs is lost in their aspectual use, as seen in the intransitive nature of the construction in (3).

Although certain varieties of aspectual meaning commonly appear in the inventory of viewpoint aspectual forms across language, such as the progressive meaning seen in (4) or the experiential meaning seen in (5), there is a high degree of idiosyncrasy in the range of aspectual meanings found in the inventory of grammaticalized viewpoint aspect in any given language. This idiosyncrasy is compounded by the tendency for aspectual meaning to be associated with meanings that are non-aspectual, so that the borderline between temporal and non-temporal meaning is not always clear. As an example, V1-*te sima(w)u*, with V2 *sima(w)u* deriving from an independent verb meaning 'put (something) away', exhibits the temporal meaning of completion of the entirety of an event up through its concluding stage, as in (7a), but also carries with it the connotation that once completed, the event cannot be undone even if one wished it so, so that the pattern comes to mark situations that are out of the control of the speaker, as in (7b), and ultimately ones that are undesirable for the speaker, as in (7c).

(7) a. *Yon-hyaku-peezi   mo   ar-u   syoosetu   o*
       four-hundred-pages extent exist-NONPST novel   ACC

       iti-niti   de yon-de   simat-ta.   (yom+te+simaw+ta)
       one-day in read-GER put.away-PST
       'He (completely) read a novel of 400 pages in one day.'

   b. *Okasi-katta kara   tui   warat-te   simat-ta. (waraw+te+simaw+ta)*
      funny-PST because inadvertently laugh-GER put.away-PST
      'It was funny so I (couldn't help) burst out laughing.'

   c. *Uti   no   ai-ken   ga   sin-de   simat-ta.   (sin+te+simaw+ta)*
      home GEN loved.dog NOM die-GER put.away-PST
      'Our beloved dog (sadly) died.'

Note how the various meanings of V1-*te sima(w)u* may be left overtly unexpressed in English. The presence of a marker such as this in Japanese, by contrast, constrains

the speaker in an important sense: while the speaker is not required to use the form, failure to do so may give rise to connotations that would be absent in English, such as, in the case of (7c), that the speaker was emotionally unaffected by the event in question.

Contrasting with viewpoint aspect are temporal meanings that are not given expression by a dedicated grammatical form, but are rather inherent to the situation or event expressed by a verb or predicate, such as whether a transition to a new state occurs in the course of the event or not, seen in (2) in the difference between *tuku* 'arrive' and *iru* 'exist'. Inherent or lexical aspect of this kind, called "situation aspect" by Smith (1997) and also known in the European tradition as "aktionsart", is covert, and therefore the temporal structure or shape in question can only be uncovered indirectly by diagnostic tests. We might expect that, being covert, such meaning would be subject to wide and idiosyncratic variation across languages. One of the major results that emerges from research on aspect, however, is that the same types of situation aspectual meaning emerge consistently across differing languages when diagnostic tests appropriate to each language are applied.

3 Zeno Vendler and Haruhiko Kindaichi

The apparently universal nature of the temporal categories making up situation aspect was demonstrated in a dramatic fashion in two classic studies of aspect carried out independently on Japanese and English in the 1950s. Both studies rely heavily on the behavior verbs exhibit with viewpoint aspectual forms as a basis for classifying them into different categories. In a famous study by Zeno Vendler (1957) of English, two sets of diagnostic tests are used to set up four categories of verbs. One is represented by the *–ing* test: whether a verb allows the *–ing* form in the progressive sense or not:

(8) a. *The children are swimming in the pool.*

  b. **The children are knowing each other.*

Another is represented by the adverbial test: which of the two temporal phrases *for X amount of time* and *in X amount of time* occurs more naturally with the verb. The predicates in (8a) and (8b) behave similarly in this regard:

(9) a. *The children swam in the pool {for 30 minutes / *in 30 minutes}.*

  b. *The children knew each other {for 2 years / *in 2 years}.*

The predicates in (10), by contrast, exhibit the opposite pattern with these adverbs:

(10) a. *John arrived at the summit {in 2 hours / *for 2 hours}.*

  b. *John mowed the lawn {in 30 minutes / ?for 30 minutes}.*

but differ from each other in terms of their acceptability with the progressive reading of *-ing:*

(11) a. *??John is arriving at the summit.* (OK in future, not progressive, sense)

  b. *John is mowing the lawn.*

These two tests (along with others) define cross-cutting classes that together form the basis for Vendler's four-way categorization of verbs into states (e.g. *know*), activities (e.g. *swim*), achievements (e.g. *reach the summit*), and accomplishments (e.g. *mow the lawn*).

Quite independently from, and substantially predating, the work of Vendler, Haruhiko Kindaichi (1950) proposed a revolutionary classification of Japanese verbs that strikingly resembles Vendler's in its methodology and results. Kindaichi's classification may be said to have laid the foundation for the subsequent study of aspect in Japanese, departing radically from earlier taxonomies that emphasized categories of verbal meaning such as transitivity and volitionality. His classification is based wholly on the behavior of verbs with the viewpoint aspectual form V-*te-iru*, specifically (a) whether or not *-te-iru* is possible in the conclusive (*shūshi*, clause-final) form of a verb, and (b) if it is possible, what the resulting interpretation is. Table 1 shows Kindaichi's four-way classification of Japanese verbs together with some representative examples. Further examples of verbs in each class are given in the discussion following Table 1.

Stative verbs (Type 1 verbs): Stative verbs do not accept the V-*te-iru* form, but occur only in the simple conclusive form (*-ru* or *-ta*) to represent a present or past state. Examples of stative verbs include *aru* 'exist (predicated of an inanimate subject)' and *dekiru* 'be able', illustrated in (12).

(12) a. *Ima, saihu no naka ni go-sen-en ar-u / *at-te-iru.*
  now wallet GEN inside in five-thousand-yen exist-NONPST/*exist-TE-IRU
  'I have 5,000 yen in my wallet right now.'

  b. *Kodomo wa eigo ga deki-ru/*deki-te-iru.*
  children TOP English NOM be.able-NONPST/be.able-TE-IRU
  'The children can (speak) English.'

In (12), the present conclusive forms *aru* and *dekiru* represent a current state or property of the subject without the aid of the aspectual form *-te-iru*. Other examples

Table 1: Four classes of verbs in the Kindaichi classification

Classification	Examples and interpretation with *-te-iru*
1. Stative verbs (Incompatible with *te iru* in the conclusive form.)	*ar u* 'be, exist' / **Heya ni tukue ga at te iru* [room LOC desk NOM exist-TE-IRU] 'There is a desk in the room'; *deki ru* 'be able' / **Kare wa eigo no kaiwa ga deki te iru* [he TOP English GEN conversation NOM be.able-TE-IRU] 'He is competent in English conversation'; *kire ru* 'cut well' / **Kono naihu wa yoku kire te iru* [this knife TOP well can.cut-TE-IRU] 'This knife cuts well'
2. Continuative verbs (V-*te iru* in the conclusive form expresses the ongoing progress of an action or event.)	*kak u* 'write' / *Kare wa tegami o kai te iru* [he TOP letter ACC write-TE-IRU] 'He is writing a letter'; *hur u* '(rain) fall' / *Ame ga hut te iru* [rain NOM fall-TE-IRU] 'It's raining'; *aruk u* 'walk' / *Kare wa toori o arui te iru* [he TOP street ACC walk-TE-IRU] 'He is walking (down) the street'
3. Instantaneous verbs (V-*te iru* in the conclusive form expresses a state resulting from the event or action of V.)	*sin u* 'die' / *Kare wa sin de iru* [he TOP die-TE-IRU] 'He is dead'; *tuk u* '(light) come on' / *Dentoo ga tui te iru* [light NOM come.on-TE-IRU] 'The light is on'; *kekkon su ru* 'get married' / *Kare wa kekkon si te iru* [he TOP get.married-TE-IRU] 'He is married'
4. Type 4 verbs (In the conclusive form, always occurs with *-te iru* to denote a characteristic attribute of the subject.)	*sobie ru* 'soar' / *Yama ga sobie te iru* [mountain NOM soar-TE-IRU] 'The mountain soars (high)'; *bakage ru* 'be foolish, stupid' / *Kimi no kangae wa bakage te iru* [you GEN idea TOP be.stupid-TE-IRU] 'Your idea is stupid'; *arihure ru* 'be common, not rare' / *Sonna hanasi wa arihure te iru* [that.kind story TOP be.commonplace-TE-IRU] 'That kind of story is common-place'

of verbs that behave in this way, qualifying them as stative, include *ataisuru* 'be worthy' (*Syoonen no yuuki-aru koodoo wa syoosan ni ataisu-ru* [boy GEN brave action TOP praise DAT be.worthy-NONPST] 'The boy's brave actions are worthy of praise') and *hanaseru* 'be compassionate/understanding (lit. can speak)' (*Syatyoo wa naka-naka hanase-ru* [boss TOP quite can.speak-NONPST] 'My boss is quite compassionate/ understanding'). As Kindaichi (1950: 50) points out, Japanese stative verbs are extremely limited in number compared with English. Verbs qualifying as stative in English commonly have counterparts in Japanese that belong to Kindaichi's instan-taneous class, such as English *have* and *know* vs. Japanese *motu* 'hold, come to have' and *siru* 'come to know' (see Section 6).

Continuative verbs (Type 2 verbs): Continuative verbs lexically express activities or events that have some duration in time without a predetermined endpoint. Unlike stative verbs, continuative verbs such as *oyogu* 'swim' in (13a) freely take *-te-iru* in the conclusive form.

(13) a. *Kodomo   wa   puuru   de   oyog-u/oyoi-de-iru.*
      children   TOP   pool   LOC   swim-NONPST/swim-TE-IRU
      'The children {will swim/are swimming} in the pool.'

    b. *Titi   wa   sibahu   o   kar-u/kat-te-iru.*
       father   TOP   lawn   ACC   mow-NONPST/mow-TE-IRU
       'Father {will mow/is mowing} the lawn.'

While the majority of verbs in this class denote volitional human activities, there are a number that represent natural events where volition is absent, such as *(Ame ga) huru* '(rain) fall' in Table 1.

Central to the Kindaichi classification is not just the compatibility or incompatibility of a verb with the auxiliary *-te-iru* but the aspectual interpretation that obtains for the resulting clause as a whole. The interpretation of the V-*te-iru* form itself falls broadly under two categories: progressive and resulting state (a variety of perfect meaning). In examples such as (13) it takes a progressive reading, indicating that the event expressed by the verb is ongoing at the time under discussion. Note that along with the *-te-iru* form, the bare non-past form is also possible in (13), with the implication that the event denoted by the verb takes place habitually or will take place in the near future, in close parallel to the contrast in English between the present tense (*He swims now/daily*) and the present progressive (*He is swimming*).

Instantaneous verbs (Type 3 verbs): Typical examples of Kindaichi's "instantaneous" verbs include *sinu* 'die', *kekkon-suru* 'get married', and most intransitive verbs that have transitive partners, such as *(doa ga) aku* '(the door) opens', *(denki ga) tuku* '(the lights) go on', and *(syukudai ga) dekiru* '(homework) becomes done'. These verbs lexically express events that are assumed to terminate in an instant. *Sinu* 'die' and *(dentoo ga) tuku* '(light) comes on', for example, express a terminal point demarcating the change from being alive to being dead, or from being off to being on. The *-te-iru* forms of these verbs do not therefore occur naturally with *saityuu da* 'be in the midst of', as in **Kare wa sin-de-iru saityuu da* [he TOP die-TE-IRU midst COP] 'He is in the midst of dying' or **Dentoo ga tui-te-iru saityuu da* [light NOM come.on-TE-IRU midst COP] 'The light is in the midst of coming on' (Kindaichi 1950: 49). Verbs like *tuku* 'arrive', *wasureru* 'forget', *usinau* 'lose', and *syuppatu-suru* 'leave' also behave in the same way.

Instantaneous verbs, like continuative verbs, can occur either in the simple non-past form or in the *-te-iru* aspectual form, as in (14).

(14) a. *Tozanka ga   yama   no   tyoozyoo ni   {mamonaku tuk-u /*
      climbers NOM mountain GEN summit   GOAL {soon   arrive-NONPST
      sudeni   tui-te-iru}.
      already arrive-TE-IRU}
      'The climbers {will arrive soon / have already arrived} at the mountain summit.'

b. *Yuusyoku wa {mainiti roku-zi ni*
dinner   TOP {everyday 6:00   at

deki-ru / moo deki-te-iru}.
be.made-NONPST / already be.made-TE-IRU}
'Dinner {is prepared at 6 o'clock every day/is already prepared}.'

As is the case with continuative verbs illustrated in (13), the simple present tense attached to instantaneous verbs denotes the happening of an event in the near future, as in (14a), or the habitual occurrence of the event, as in (14b). Unlike with continuative verbs, however, *-te-iru* with instantaneous verbs conveys a perfect meaning, indicating a state that results from the prior occurrence of the event expressed by the verb. Thus *tui-te-iru* in (14a) implies that the climbers now stand on the summit of the mountain, and *deki-te-iru* in (14b) that dinner is ready to eat. This is because *-te-iru* focuses on the time after the event of an instantaneous verb has occurred (see Section 4). Note in this connection that the verb *dekiru* has two uses, one as a stative verb denoting a person's ability, seen earlier in (12b), and the other as an instantaneous verb meaning 'become complete, come into existence', illustrated in (14b). This can be seen as a case of homophony, two verbs pronounced identically belonging to different aspectual classes.

A similar duality of aspectual behavior can be seen with such verbs as *mieru* 'look like', *wakaru* 'understand', and *hukumu* 'contain', which behave like stative verbs in Kindaichi's classification in their ability to denote a present state in their bare non-past form. Curiously, however, these verbs can also take the *-te-iru* form to represent a present state – a peculiar behavior Kindaichi failed to observe. Also exhibiting this behavior are stative predicates belonging to the class of Sino-Japanese verbal nouns, such as *zoku-suru* 'belong', *sonzai-suru* 'exist', and *imi-suru* 'mean'.

(15) *Kuzira wa   honyuurui   ni   zoku-su-ru/zoku-si-te-iru.*
whale TOP mammal.class DAT belong-NONPST/belong-TE-IRU
'Whales belong to the mammal class.'

Unlike the polysemous verb *dekiru*, however, these verbal nouns cannot be treated as homophonous pairs, one member of the pair having membership in the stative class and the other in the instantaneous class, as there is no perceptible semantic difference in meaning between their bare form and *-te-iru* form. This suggests the need for a finer-grained classification, a topic taken up in Section 5.

The key insight advanced by Kindaichi in the above was to attribute the difference between the progressive and perfect (resulting state) interpretations of V-*te-iru* to the lexical character of the attached verb. Verbs that yield the progressive meaning with V-*te-iru* (e.g. *oyogu* 'swim') are continuative verbs, and those that yield the perfect (resulting state) meaning (e.g. *(tyoozyoo ni) tuku* 'arrive (at the summit)') are instantaneous verbs. The "instantaneous" label is clearly motivated by the sense that events expressed by such verbs have no extension in time, and therefore cannot at the time of speech be seen as ongoing.

Type 4 verbs: There is a fourth, much smaller category, of verbs that exhibits yet a different sort of behavior with V-*te-iru* that Kindaichi called simply "Type 4" verbs. These are verbs that *must* occur in the V-*te-iru* form when occurring in finite, sentence-final contexts, as illustrated in (16).

(16) *Kare   no   gengo-nooryoku   wa   sugure-te-iru / *sugure-ru.*
he   GEN   linguistic-ability TOP   excel-TE-IRU / excel-NONPST
'His linguistic ability is outstanding.'

These verbs are like instantaneous verbs in having a V-*te-iru* form that receives a stative interpretation, but are unlike them in lacking a bare non-past -*ru* form. Essentially, these verbs preserve the form and meaning of a perfect (resulting state) construction, but have had the occurrence of an originating event "bleached" from their meaning structure, leaving only the stative part of the structure. The close affinity between Type 4 and instantaneous verbs can be seen in certain examples of V-*te-iru* forms where either behavior is possible, as in (17), depending on whether an event giving rise to the stative meaning can be conceptualized or not.

(17) a. *Kugi   ga   magat-te-iru / magar-u.*
nail   NOM   bend(intr.)-TE-IRU / bend-NONPST
'The nail {is bent / will bend}.'

   b. *Miti   ga   magat-te-iru / *magar-u.*
road   NOM   bend(intr.)-TE-IRU / bend-NONPST
'The road {is bent (curves) / will bend}.'

In this sense, Kindaichi's Type 4 class may be seen as a derivative of the instantaneous category. Other examples of verbs in this class include *ni-te-iru* 'resemble', *(yama ga) sobie-te-iru* '(the mountain) soars (high)', and *arihure-te-iru* 'be commonplace'.

Leaving aside the Type 4 class, then, a correspondence emerges between the three major categories of Kindaichi's classification and counterparts to them in Vendler's classification, as shown in Table 2.

Table 2: Corresponding categories of Vendler and Kindaichi

Vendler	Kindaichi
States	Stative verbs
Activities	Continuative verbs
Achievements	Instantaneous verbs

These correspondences are borne out by numerous grammatical parallels, of which three must suffice here as examples. One is the parallel, seen earlier, between the failure of states to allow the *-ing* form in English and the failure of stative verbs in Japanese to allow the V-*te-iru* form. A second is the distinction seen in both languages in the tense interpretation of stative verbs, which in their bare form are able to denote a situation that is true at the moment of speech, as opposed to the other two categories, which in their bare form denote instead either a future or habitual occurrence. This is illustrated in (18) in the tense interpretation of the stative verb *aru* 'exist' versus the instantaneous (achievement) verb *butukaru* 'bump into'.

(18) a. *Teeburu no   ue   ni   hon   ga   san-satu   ar-u.*
        table   GEN top LOC books NOM three-CLF exist-NONPST
        'There are three books on the table.'

   b. *Abuna-i!*   *Mae   no   kuruma ni   butukar-u*   *yo.*
      be.dangerous-NONPST front GEN car   GOAL bump.into-NONPST SFP
      'Watch out! You('ll) bump into the car in front of you'

Stative verbs and activity (continuative) verbs, by contrast, behave like each other, and unlike achievement (instantaneous) verbs, in their ability to co-occur with temporal adverbial phrases in *X-(kan)* 'for (X amount of time)' and *X de* 'in (X amount of time)', as illustrated in Table 3.

Table 3: Compatibility with temporal adverbials

	sanzikan de 'in 3 hours'	*sanzikan* 'for 3 hours'	*sono syunkan ni* 'at that moment'
States	...**soko ni ita* 'was there'	...*soko ni ita*	...*soko ni ita*
Activities (continuative)	...**hataraita* 'worked' ...**hasitta* 'ran'	...*hataraita* ...*hasitta*	...**hataraita* ...**hasitta*
Achievements (instantaneous)	...*Kyooto ni tuita* 'arrived in Kyoto' ...*sinda* 'died'	...**Kyooto ni tuita* ...**sinda*	...*Kyooto ni tuita* ...*sinda*

Stative verbs act like achievements, on the other hand, and unlike activities, in their ability to co-occur with temporal adverbs indicating instants of time such as *sono syunkan ni* 'at that moment' in Table 3.

A category of Vendler's that is prominently absent in Kindaichi's classification is that of accomplishments (*mow the lawn, write a novel*, etc.). As seen earlier in (10) and (11), accomplishments in English are like activities in allowing the *-ing* form, but are like achievements in occurring more comfortably with temporal phrases in *in* rather than *for*. Though overtly absent in Kindaichi's classification, accomplishments covertly make their presence felt in a class of examples Kindaichi noted as

unusual in allowing either a progressive or a perfect interpretation with V-*te-iru*, as in (19). Kindaichi treated these predicates as alternating in their membership, sometimes belonging to the activity (continuative) and sometimes to the achievement (instantaneous) class, depending on which of the two interpretations they receive with the V-*te-iru* form.

(19) a. *Kare wa syoosetu o kai-te-iru.*
      he   TOP novel   ACC write-TE-IRU
      'He {is writing/has written} novel(s).'

    b. *Hanako wa kimono o ki-te-iru.*
      Hanako TOP kimono ACC put.on-TE-IRU
      'Hanako {is putting on/is wearing (has put on)} a kimono.'

Whether a predicate exhibits this kind of alternating behavior is not, however, something that follows from its lexical meaning alone, but from contributions to its meaning made by direct objects and other nouns in the same clause that have the effect of "delimiting" the event expressed (Tenny 1994). The delimiting effect of noun phrases on the aspectual character of clauses as a whole has been noted for English by Dowty (1979) and others, who point out that a given verb may exhibit either an accomplishment or an activity reading depending on whether its direct object is a definite noun phrase, as in (20a), or a mass or plural noun phrase, as in (20b).

(20) a. *He wrote that mystery novel in three months.*

    b. *He wrote mystery novels for his whole life.*

In the case of (19a) and (20a), the event of writing a novel consists of an activity of writing, but one that continues only up to the point that the novel reaches its conclusion, and ceases after that. The endpoint of the writing itself has the character of an achievement in that it has no extension in time – i.e., is instantaneous in Kindaichi's sense, forming a boundary point between the activity of writing and its absence. The meaning of an accomplishment can therefore be seen to be decomposable into other categories of lexical aspect, and is in that sense secondary to, the three basic categories of states, activities, and achievements that comprise the fundamental "building blocks" of lexical aspect in both English and Japanese in the Vendler and Kindaichi frameworks. Furthermore, since accomplishment is a category of meaning to which elements in a clause outside of the predicate contribute, it is not strictly speaking a category of meaning inherent to the predicate itself, and does not in that sense belong to the category of lexical aspect. Still, it is the predicate that fundamentally determines the pattern of non-predicate elements that may

occur in a clause, primarily by means of its argument structure, so the lexical meaning of a predicate is an indispensable component of accomplishment meaning even at the clausal level.

4 Homogeneous and non-homogeneous aspectual meaning

Of the three fundamental categories of situation (lexical) aspect, states and activities are distinct from achievements in consisting of situations that are unchanged, or "homogeneous," over some interval, as reflected in their ability to co-occur with *for* phrases in English and zero-marked phrases such as *san-zikan* 'for 3 hours' in Japanese (see Table 3). But states and activities also exhibit important differences. States, for example, can be predicated of an instant of time in a way that activities cannot: in the non-past bare form, as in (21a), a stative predicate receives a literal present interpretation, whereas an activity verb receives a future interpretation, and in past tense contexts, as in (21b), a state predicated of a moment of time is again seen to obtain at that moment whereas an activity is not.

(21) a. *Kare wa ima soko ni i-ru.*   VS *Kare wa ima hasir-u.*
      he TOP now there LOC be-NONPST   he TOP now run-NONPST
      'He is there now.'   'He now (will) run.'

   b. *Kare wa sanzi ni soko ni i-ta.*   VS *Kare wa sanzi ni hasit-ta.*
      he TOP 3:00 at there LOC be-PST he TOP 3:00 at run-PST
      'He was there at 3:00.'   'He ran at 3:00.'

The moment of 3:00 with the past form of the activity *hasit-ta* in (21b) is most naturally interpreted as the point from which the activity of running (or some activity preparatory to that) begins, but not a point in time at which the activity of running itself obtains, a meaning that would require the use of the distinct past progressive form *hasit-te-ita* 'was running'.

    The reason for this difference between states and activities becomes apparent when we apply a test for homogeneous meaning called the "subinterval test" (Taylor 1977; Dowty 1979). States are purely homogeneous in that a state predicated over an interval of time can also be predicated of any subinterval of that time, no matter how small, even down to an instant. So if it is the case that someone was at a certain place between 2:00 and 4:00, it is possible to say that s/he was at that place at any moment within that interval, such as the moment of 3:00. Activities, by contrast, while they may be seen to be homogeneous for the purpose of intervals of time of some length, are not fine-grained enough in their homogeneity to be predicated of

either instants or very small intervals of time. So while it is possible to say that someone ran between 2:00 and 4:00, it is not possible to say that s/he ran at 3:00 (the facts are the same here for English and Japanese). This is because activities are composed of the repeated occurrence of what Dowty (1979) calls "subroutines" (also "atoms" of meaning in the framework of Bohnemeyer and Swift 2004), that themselves have extension in time larger than instants. The activity of running, for example, is made up of the repetition of a subroutine extending from the point one foot is raised off of the ground through the point that it once again makes contact with the ground and until it once again is raised off of the ground. Such subroutines are too large to be predicated of instants of time, although when viewed over a large enough interval of time, can together be construed as roughly homogeneous, much like the impression of linear smoothness that results when viewing a dense series of dots through a camera zoom lens adjusted to a wide angle. The question may arise why then activities can be predicated of instants of time through use of the progressive V-*te-iru* form, as in *3-zi ni hasit-te-ita* 'was running at 3:00'. This is because the V-*te-iru* form, itself stative-like in being predicable of either an interval or instant of time, imposes a perspective that views that interval or instant as being properly contained in a larger interval over which the situation expressed by the verb is seen to occur, thereby escaping the conflict that would arise in the bare form (e.g. *3-zi ni hasit-ta* 'ran at 3:00') between instantaneous time reference and an event type that cannot be seen as obtaining instantaneously.

The difference in the temporal structure of states and activities can be schematized as in (22), with states represented as a continuous, purely homogeneous line and activities as a repetition of subroutines, individually symbolized as ~, which in the aggregate take on a homogeneous character as the cognitive "zoom lens" takes in larger intervals of time (the schematization here is a modification of that adopted in Shirai 2000).

(22)   States: ———————————   (e.g. *ar-u* 'exist')

      Activities: ~~~~~~~~~~~~~~~~~~   (e.g. *hasir-u* 'run')

The non-homogeneous character of achievements, by contrast, can be schematized as a unique boundary, symbolized as (X) in (23), that separates two distinct states Q and ¬ Q ('not Q'), or vice versa, existing in a relationship of affirmative and negative. Taking the case of the achievement verb *sinu* 'die', for example, Q corresponds to the state of being dead and ¬ Q of not being dead.

(23)   Achievements: ——— ¬ Q ——— (X) ——— Q ———   (e.g. *sin-u* 'die')

Accomplishments, then, combine the temporal character of activities and achievements as in (24), having essentially the same shape as (23) except that the boundary

(X) is preceded by the absence of a state that coincides with the presence of an activity and is followed by the presence of a new state that coincides with the absence of that activity. Taking *syoosetu o kaku* 'write a novel', as an example, Q is a state characterized by the existence of a completed novel and ¬ Q a state characterized by its non-existence.

(24)   Accomplishments: ~~~ ¬ Q ~~~ (X) ——— Q ———   (*syoosetu o kak-u* 'write a novel') (=Activity + Achievement)

Within the Kindaichi framework, (X) is understood as an instantaneous moment, but the question can be raised as to whether the boundary between ¬ Q and Q is indeed instantaneous for all achievements. The question also arises as to whether the three fundamental categories of situation aspect underlying the Vendler/Kindaichi model are sufficient to describe exhaustively all types of situation aspect. Both these questions are raised in an influential critique of the Kindaichi framework by Okuda (1978).

5 Okuda's criticism of Kindaichi: degree achievements, semelfactives, and transitivity pairs

In his critique of Kindaichi (1950), Okuda (1978) argues that an approach to Japanese aspect that seeks to account for the temporal behavior of predicates in quantitative terms such as those implicit in the terms "instantaneous" and "continuative" is fundamentally misguided. He points out that there are, on the one hand, verbs as in (25a) that receive a perfect interpretation in the V-*te-iru* form but which are not understood to express instantaneous occurrence, and, on the other, verbs as in (25b) that appear to express instantaneous occurrence but do not receive a perfect interpretation in the V-*te-iru* form.[2]

(25)   a.   *hutor-u* 'become fat' → *hutot-te-iru* 'is (in the state of having become) fat'
tukare-ru 'become tired' → *tukare-te-iru* 'is (in the state of having become) tired'

   b.   *mabatak-u* 'blink' → *mabatai-te-iru* 'is blinking'
hiramek-u 'flash' → *hiramei-te-iru* 'is flashing'

---

2 Kindaichi was not unaware of the existence of exceptional verbs of these two types. He points out very briefly in the final section of Kindaichi (1950: 62) examples of verbs like those in (25a) such as *nobiru* 'lengthen, get longer' and *nigoru* 'get muddy' and like those in (25b) such as *mabataku* 'blink' and *bekken suru* 'catch a glimpse', but did not pursue the implications these verbs posed for his general classification.

The examples in (25a) anticipate what Dowty (1979) was to call "degree achievements", verbs that express a boundary between the presence or absence of a state, but one that can be seen to obtain in varying degrees, making it difficult to identify an instant in time at which the final state may be said to have arrived. Representative examples of such verbs include *cool down* (Japanese *sameru* [intr.]) and *warm up* (Japanese *atatamaru* [intr.]). The events expressed here are more naturally seen to occur over an interval of time, at the lower (temporal) bound of which the state in question obtains to an incipient extent and at the upper bound of which the state obtains to its full extent. What counts as the lower and upper bounds will of course depend on the conversational purpose at hand. In the case of becoming fat, a weight that would count as the upper bound for a 7-year old girl will be lower than that for a 45-year old adult man; indeed, the higher bound in the former case is likely to be lower than the lower bound in the latter case. While advocates of the Kindaichi approach might argue that there is in all such cases an instant, however arbitrary and imagined, that acts as a divide between the presence and absence of the relevant state, it seems in better accord with native speaker intuition to conceptualize the boundary (X) in (23) as a bounded *interval* having some extension in time, one that approaches an instant as a special case as the upper and lower bounds merge and ultimately overlap.

Turning to the examples in (25b), these correspond to predicates that have been called "semelfactive" (a term attributed to Comrie (1976)). Besides those in (25b), this category includes verbs like *knock* (Japanese *tataku*), *sneeze* (Japanese *kusyami o suru*), *cough* (Japanese *seki o suru*), and others that express events occurring over a very short interval of time, so short as to sometimes approach instants. Such events are not of course literally instantaneous, in the sense of having no extension in time, as a slow motion replay of any semelfactive event captured on film would clearly reveal. Rather, they may be identified with the individual "atomic" subroutines that make up activities, symbolized by ~ in the schema in (22), and as such each has a distinct beginning, end, and middle. The critical difference between semelfactive events and subroutines of activities is in the way the latter occur repeatedly. Semelfactive events, by contrast, occur as unique, stand-alone events, and are in that sense similar in temporal structure to achievements, as schematized in (23), with the difference that (X) does not form a boundary between two distinct states as it does with achievements.

(26)   Semelfactives: —— Q —— (X) —— Q ——   (e.g. *tatak-u* 'knock')

The affinity of semelfactives to activities nevertheless becomes clear in the case of iterated occurrence. As seen in (25b), the iterated occurrence of semelfactives licenses a progressive interpretation with V-*te-iru* in Japanese, just as does the -*ing* form in corresponding cases in English, in a way that makes the interpretation indistinguishable from that given to garden-variety activity verbs.

Although semelfactive predicates do not technically represent instantaneous occurrence, neither are they "continuative", in Kindaichi's sense, and so their progressive interpretation with the V-*te-iru* form is left unaccounted for in the Kindaichi framework. Semelfactives, together with non-instantaneous degree achievements that nevertheless receive a perfect interpretation with V-*te-iru*, thus point to a fundamental difficulty with accounting for aspectual categories in purely quantitative terms defined by length of time of occurrence such as "continuative" and "instantaneous". Instead, what is critical for Okuda in determining the aspectual behavior of the V-*te-iru* form is whether V expresses an event that involves a *change of state* of some kind in the "subject" of the clause it appears in. Okuda's most central insight for the analysis of aspectual meaning consists in the evidence he brings to bear on this from the behavior of transitive/intransitive verb pairs.

Transitive and intransitive (sometimes referred to as causative and inchoative) verbs in Japanese typically fall into morphologically related pairs with the subject of the intransitive member corresponding to the direct object of the transitive member as in (27), with examples given in (28).

(27)   NP$_1$ *ga* NP$_2$ *o* V$_{tr}$ $\leftrightarrow$ NP$_2$ *ga* V$_{in}$

(28) a. *Taroo ga   mado   o   ake-ta.* $\leftrightarrow$ *Mado   ga   ai-ta.*
        Taro   NOM window ACC open$_{tr}$-PST   window NOM open$_{in}$-PST
        'Taro opened the window.'   'The window opened.'

     b. *Tokeiya   ga   tokei o   naosi-ta.* $\leftrightarrow$ *Tokei ga   naot-ta.*
        clock.repairman NOM clock ACC fix$_{tr}$-PST   clock NOM fix$_{in}$-PST
        'The clock repairman fixed the clock.'   'The clock got fixed.'

Okuda noted that verbs falling into such pairs receive differing default interpretations with the V-*te-iru* form, transitive verbs receiving a progressive interpretation and intransitives a perfect (resulting state) interpretation.

(29) a. *Taroo ga   mado   o   ake-te-iru.* $\leftrightarrow$ *Mado   ga   ai-te-iru.*
        Taro   NOM window ACC open$_{tr}$-TE-IRU   window NOM open$_{in}$-TE-IRU
        'Taro is opening the window.'   'The window is open.'

     b. *Tokeiya   ga   tokei o   naosi-te-iru.* $\leftrightarrow$ *Tokei ga   naot-te-iru.*
        clock.repairman NOM clock ACC fix$_{tr}$-TE-IRU   clock NOM fix$_{in}$-TE-IRU
        'The clock repairman is fixing the clock.'   'The clock is fixed.'

For Okuda, who viewed corresponding transitive and intransitive clauses as different manifestations of a single event involving no difference in the length of time required to occur, this behavior with V-*te-iru* would be impossible to account for in the Kindaichi framework, where the intransitive verb expresses an "instantaneous", and the transitive verb a "continuative" event. Instead, he argued, the critical differ-

ence between the two constructions is accounted for by the grammatical role played by the noun phrase seen to undergo a change in state (the "theme"), represented by NP_2 in (27). In the intransitive case, the theme NP_2 appears prominently as subject, highlighting change in state as a characteristic feature of the meaning of the intransitive clause. In the transitive case, the subject represents an entity seen to perform an activity, not undergoing a change in state, thereby licensing, for Okuda, the progressive reading.

Although transitive/intransitive pairs provide convincing evidence for the role of change of state in the aspectual behavior of predicates, Okuda's assumption that corresponding transitive and intransitive constructions express identical events with identical temporal structures cannot be maintained in light of the fact that transitive constructions include in their meaning structure an activity component that is absent in intransitive constructions. The activity in a transitive construction, however, is seen to cease at the point in time that the object entity takes on a new state. In (28a) above, for example, the activity of Taro opening the window ceases at the point the window becomes open. This means that transitive constructions have in essence the temporal shape of accomplishments (as in (24), repeated in (30)), in contrast to intransitive constructions, which have the temporal structure of achievements (as in (23), repeated in (31)). The boldface here indicates the portion of temporal meaning that is literally denoted by the lexical transitive and intransitive verb constructions, respectively, with the non-boldface indicating portions of temporal meaning that form a background to this.

(30) Transitive: **~~~ ¬ Q~~~(X)** —— Q —— (e.g. *mado o ake-ru* 'open window')

(31) Intransitive: —— ¬ Q —— **(X)** —— Q —— (e.g. *mado ga ak-u* 'window opens')

If transitive constructions encode an accomplishment structure, then it should be possible for them to exhibit not only progressive, but also perfect, interpretations with the *-te-iru* form, and this is in fact the case. The sentences in (32) are examples of transitive constructions that receive such non-progressive, perfect meaning.

(32) a. *Taroo wa heya ni i-ru toki itumo mado o ake-te-iru.*
Taroo TOP room LOC be-NONPST when always window ACC open$_{tr}$-TE-IRU
'When in his room, Taroo always has his window open
(lit. has opened his window).'

b. *Tokeiya wa nandomo kono tokei o naosi-te-iru.*
clock.repairman TOP many.times this clock ACC fix$_{tr}$-TE-IRU.
'The clock repairman has fixed this clock many times.'

These examples in turn call into question a second assumption made by Okuda – that it is a change of state in the *subject* that is the primary determining factor for

the aspectual behavior of V-*te-iru*. The presence of a thematic argument (NP$_2$ in (27)) is sufficient to license a perfect interpretation with V-*te-iru* regardless of whether it occurs in subject position or direct object position, as long as it constitutes an "internal argument" in the sense of Williams (1981). In terms of the schemata in (22)–(24), a necessary condition for assigning a perfect interpretation to V-*te-iru* is the presence in aspectual structure of a boundary (X) between two *distinct* states or between the occurrence and non-occurrence of an activity, regardless of how the boundary is introduced, whether inherent to the lexical aspectual meaning of the predicate itself, or introduced by elements co-clausal with the predicate such as internal arguments, goal phrases, and adverbial phrases that delimit in some sense the temporal extent of the event denoted by the predicate.

The possibility of two interpretations for V-*te-iru* with transitive verbs (e.g. *Mado o ake-te-iru* [window ACC open$_{tr}$-TE-IRU] 'is opening/has opened the window') but only one with intransitive verbs (e.g. *Mado ga ai-te-iru* [window NOM open$_{in}$-TE-IRU] 'the window is (has become) open', but not 'the window is opening (is changing from closed to open)') falls out naturally from the accomplishment vs. achievement distinction inherent in transitive vs. intransitive constructions in combination with an aspectual "viewpoint" contributed by *-te-iru* itself that is imposed on each of these.[3] This aspectual viewpoint may be characterized as having two components. One is to present an aspectually homogeneous situation, that is, a situation possessing the "subinterval property", as obtaining over an interval *surrounding* the time under discussion (called "topic time" in Klein 1994 and "time of reference" in Reichenbach 1947). As seen earlier, the subinterval property characterizes both states and activities, so this interval may be seen as consisting either of a state (as in (33a)) or an activity (as in (33b)).

(33)   Aspectual structure of *-te-iru* (T = topic time):

    a.   T
        [——————————————————] (e.g. *sonzai-si-te-iru* 'exists')[4]

    b.   T
        [~~~~~~~~~~~~~~~~~~~~~~~~] (e.g. *hasit-te-iru* 'is running')

---

3 The analysis of the aspectual meaning of V *te iru* presented here develops and expands on ideas presented in an earlier form in Jacobsen (1984, 1992) and summarized in Tsujimura (1996).

4 Kindaichi defined stative verbs as verbs that do not allow *te iru*, so examples like (33a) should not in theory occur. This follows from the fact that stative verbs inherently possess a homogeneous aspectual character, so that there is no need to redundantly impose on them the homogeneous aspectual viewpoint contributed by *te iru*. The possibility of V *te iru* forms that express a pure state (as opposed to a resulting state) nevertheless arises in the case of "Type 4" verbs (see examples (16) and (17)) and of certain verbs that alternate between the stative and Type 4 categories. These are infrequent among native Japanese verbs, but are not uncommon among Sino Japanese verbs such as *sonzai suru* 'exist' (see Section 6 for more examples).

While activities and states are both homogeneous in their aspectual structure, achievements and semelfactives are not, containing as they do a uniquely occurring (X) in their aspectual structure (see (23) and (26)). Any interval containing, or consisting of, this (X) will contain subintervals that are distinct from (X) itself, and thus fail the subinterval test defining homogeneous predicates. This means that the aspectually homogeneous interval defined by V-*te-iru* cannot include within itself any occurrence of such a unique (X). When V in V-*te-iru* is itself an achievement predicate, then, the aspectual structure of achievements in (22), repeated here as (34), will dictate that the interval defined by V-*te-iru* can only coincide with the homogeneous subcomponents of the structure, those corresponding to either the state $\neg$ Q preceding (X) or the state Q following it. With an achievement verb such as *aku* 'open$_{in}$', for example, Q represents the state of being open and $\neg$ Q the state of not being open.

(34)   Achievements: —— $\neg$ Q —— (X) —— Q —— (e.g. *ak-u* 'open$_{in}$')

Of these two possibilities, however, only the second arises in Japanese, because of a second component to the viewpoint aspectual meaning imposed by -*te-iru*. As noted by Soga (1983), McClure (1995), and others, V-*te-iru* presents the situation expressed by the predicate V as being *actually* realized, either in whole or in part, a feature probably related to the historical genesis of the gerund form *te* from a marker of perfect (completive) aspect. This means that in (34), the situation expressed by the predicate, in this case (X) itself, must be realized *prior to* topic time. This yields a perfect interpretation on a form such as *ai-te-iru* 'be open (in a state of having become open)' where the event of opening must precede the homogeneous interval surrounding topic time. The meaning imposed by -*te-iru* on achievement verbs can be represented as in the bracketed portion of the schema in (35).

(35)   Interpretation of -*te-iru* with achievements (T = topic time):

$$\text{—— } \neg \text{ Q —— (X) —— [—— } \underset{}{\overset{\text{T}}{\text{Q}}} \text{ ——] } -$$

(e.g. *ai-te-iru* 'be open')

Accomplishments, by contrast, having the aspectual structure in (24), repeated here as (36), offer two distinct sites where the homogeneous interval defined by V-*te-iru* may be situated, one the state following (X), as in the case of achievements, the other the activity interval preceding (X).

(36)   Accomplishments: ~~~ $\neg$ Q ~~~ (X) —— Q ——
       (*mado o ake-ru* 'open a window') (=Activity + Achievement)

Which of these two sites is selected will yield alternately a perfect or a progressive reading in a V-*te-iru* form constructed from an accomplishment, as in *mado o ake-te-iru* 'have opened/be opening a window'. These two possibilities are represented by the bracketed portions of the schemata in (37a) and (37b).

(37) Interpretation of *-te-iru* with accomplishments (T = topic time):

    a.
$$\sim\sim\;\neg\;Q\sim\sim\;(X)\;\overset{\textstyle T}{\longrightarrow}\;[\overline{\quad Q \quad}]\;—$$
    (e.g. *mado o ake-te-iru* 'has opened the window')

    b.
$$\sim\sim[\sim\sim\;\neg\;Q\sim\sim]\sim\sim(X)\;\overset{\textstyle T}{\longrightarrow}\;Q\;\overline{\quad\quad}$$
    (e.g. *mado o ake-te-iru* 'is opening the window')

Because of the composite character of accomplishments, formed from the combination of an activity and an achievement, either of these readings satisfies the requirement that *-te-iru* present a situation as actually realized: in the case of (37a), the achievement component (X) has been realized prior to topic time, and in the case of (37b), a portion of the activity component has been realized prior to topic time.

    While Okuda's focus on the role of the subject as either undergoing a state or performing an activity thus fails to capture with sufficient generality the lexical aspectual structures responsible for the behavior of the V-*te-iru* form, it remains the case that of the two possible interpretations of V-*te-iru* with transitive constructions, the progressive reading is the less marked, default reading, providing some support for Okuda's claim that change of state (as encoded in the (X) component of the structure in (30)) is less prominent in transitive constructions than intransitive constructions. More importantly, his focus on the differing aspectual behavior of transitive and intransitive patterns is to be credited with bringing to light the central role of change of state as a qualitative feature determining the aspectual behavior of Japanese predicates, providing a corrective to the exclusively quantitative approach of Kindaichi that views length of time of occurrence of an event as determinative of such aspectual behavior.

    When considered from a cognitive perspective, however, qualitative dimensions of aspect necessarily have quantitative dimensions as well, so that approaches highlighting one or the other, are in fact complementary to, rather than incompatible with, one another. Despite the existence of degree achievements considered earlier, the fact is that typical change of state achievements such as *hazimaru* 'start', *owaru* 'finish', *tuku* 'arrive', *hairu* 'enter', and *kowareru* 'become broken', for example, focus purely on the realization of a new state, without any reference to a process by which the state is brought about, and are therefore seen as occurring instantaneously. Activities, by contrast, whether simple such as *hasiru* 'run', *odoru* 'dance', and *oyogu* 'swim', or incorporated as a component of meaning in accomplishments,

such as *syoosetu o kaku* 'write a novel', *mado o akeru* 'open a window', and *tokei o naosu* 'fix a clock', are typically seen as sustained by an agentive force internal to the human subject, which intends some motion or change of state in the outside world (Searle 1983). Since it is impossible for the intention and the resulting motion or change to occur absolutely simultaneously – the intention in the mind must be transmitted to the real world through neurons and muscles inside the body and, in some cases, tools outside the body – activities and other agentive actions will necessarily have a quantitatively durative character to them (Jacobsen 1997). Qualitative and quantitative dimensions of lexical aspectual structure are, in this way, mutually interrelated and, in the final analysis, no approach that focuses on one to the exclusion of the other can be seen to provide a satisfactory account of aspectual meaning in Japanese.

6 Morpho-syntactic categories and lexical aspect in Japanese

The temporal behavior of predicates in both English and Japanese considered in the previous sections points convincingly to the existence of a common core of situational (lexical) aspectual categories underlying the two languages, and indeed all natural languages. While the categories themselves may be universal, however, numerous idiosyncrasies exist between Japanese and English in the category membership of corresponding lexical items in the two languages.

Unlike English, Japanese has three predicate types that inflect for tense – verbs, adjectives (identified by their *–i* ending, as in *taka-i* 'be high', *samu-i* 'be cold', etc.) and copular expressions (occurring with the inflecting copula *da*, including both 'adjectival nouns' such as *kirei da* 'be pretty', *sizuka da* 'be quiet', etc. and noun + copula constructions, such as *gakusei da* 'be a student', *byooki da* 'be (in a state of) sickness', *hantai da* 'be (in an attitude of) opposition', etc.). Adjectives and copular expressions are inherently stative in meaning and make up the vast majority of the membership of the class of stative predicates. The number of strictly stative verbs – ones that do not allow the V-*te-iru* form–is, by contrast, very small, limited among native verbs to *aru* 'exist (inanimate)', *iru* 'need', *dekiru* 'be able (to do)', *wakaru* 'understand', and *iru* 'exist (animate)'. There are additionally a handful of native verbs that are stative either in the bare form or the V-*te-iru* form, such as *ataru/ atat-te-i-ru* 'correspond to', *a(w)u/at-te-iru* 'match', etc., and some Sino-Japanese and onomatopoetic expressions formed on *suru* capable of a stative interpretation, but which likewise allow either the bare or V-*te-iru* form under that interpretation: *zoku-suru/si-te-iru* 'belong', *sonzai-suru/si-te-iru* 'exist', *husoku-suru/si-te-iru* 'be lacking', *huwahuwa-suru/si-te-iru* 'be fluffy', *hinyari-suru/si-te-iru* 'be chilly', etc. Among the few stative verbs in the native lexicon, furthermore, *iru* 'exist (animate)'

and *wakaru* 'understand' exhibit numerous non-stative characteristics, such as allowing the imperative, causative, and (in the case of *iru* 'exist (animate)'), passive forms. The number of verbs whose lexical meaning is strictly stative in Japanese is therefore very limited.

Of the relatively larger inventory of stative verbs in English, the closest lexical counterparts in Japanese are often members of other morphosyntactic categories, such as adjectives (e.g. *hosii* 'want') copular expressions (e.g. *suki da* 'like (find attractive)'), or other verbal lexical categories – e.g. the examples in (38), where the stative meaning is expressed by means of change-of-state achievement verbs to which V-*te-iru* has been attached, yielding a perfect (resulting state) interpretation

(38)   a.   *Sir-u* 'come to know' → *Sit-te-iru* 'know' (lit. 'have come to know')

  b.   *Mot-u* 'hold, come to have, acquire' → *Mot-te-iru* 'have, own'
       (lit. 'be in the state of having acquired')

Cross-over between adjective and verb categories is also frequently seen in the opposite direction, with meanings expressed adjectivally in English often expressed through the V-*te-iru* form of change-of-state achievement verbs in Japanese, such as in (39).

(39)   a.   *ak-u* 'open$_{in}$' → *ai-te-iru* '(be) open' (lit. 'be in the state of having
       become open', similarly in examples following)

  b.   *kawak-u* 'dry$_{in}$ (become dry)' → *kawai-te-iru* '(be) dry'

  c.   *nure-ru* 'get/become wet' → *nure-te-iru* '(be) wet'

  d.   *hutor-u* 'get/become fat' → *hutot-te-iru* '(be) fat'

In contrast to the small number of stative verbs in Japanese, the number of achievement verbs in its verbal lexicon is disproportionately large. As a result, Japanese exhibits a more consistent correlation between morphological category and aspectual meaning than does English: lexical stative meaning is concentrated in adjectival and copular forms, while verbs are almost exclusively non-stative in their lexical meaning, stative interpretations of verbs largely being limited to the V-*te-iru* form.

Recall from the discussion of Kindaichi's Type 4 verbs in Section 3 that V-*te-iru* forms built on achievement verbs do not always require that the event represented by the achievement verb have actually occurred. This is true also of the verbs in (38) and (39), as seen in example (40a), where the state of a hole existing could not have arisen from any prior event of a hole forming. Contrast this with the explicit denotation of the occurrence of the event itself when the achievement verb occurs in its bare form, absent *te-iru*, as in the past tense example in (40b).

(40) a. *Nihon no gozyuu-en-dama ni wa ana ga ai-te-iru.*
 Japan GEN 50-yen-coin LOC TOP hole NOM become.open-TE-IRU
 'Japanese 50-yen coins have a hole in them.'

 b. *Ansyoo ni noriage-te hune no soko ni ana ga ai-ta.*
 reef LOC run.aground-CONT boat GEN bottom LOC hole NOM open$_{in}$-PST
 '(We) ran aground on a reef and a hole was formed in the bottom of the boat.'

Intermediate to (40a) and (40b) are uses of the V-*te-iru* form that either involve explicit reference to the prior occurrence of an event, as in the "historical record" use of V-*te-iru* in (41a), or where such an event can be imagined, if not explicitly referred to, as in (41b).

(41) a. *Zyuu-nen mae no oozisin de mati no hiroba ni*
 10-year ago GEN big.quake INS town GEN square LOC

 ookina ana ga ai-te-iru.
 large hole NOM open$_{in}$-TE-IRU
 'A large hole (is on record as having) opened up in the town square in the big earthquake 10 years ago.'

 b. *(Miru to) hune no soko ni ookina ana ga ai-te-ita.*
 look COND boat GEN bottom LOC large hole NOM open$_{in}$-TE-ITA
 'When I looked, there was a large hole in the bottom of the boat.'

The examples in (40) and (41) exhibit a shift in the portion of the aspectual structure denoted by the verb, the portion indicated in boldface in (42), from the "standard" denotation of an achievement event (X) in (42a), to a larger swath of the structure including the resulting state in (42b), to a defocusing of the achievement event (X) in (42c), and ultimately to a total exclusion of such an event in (42d).

(42) a. —— ¬ Q —— **(X)** —— Q —— (e.g. (40b))
 b. —— ¬ Q —— **(X)** —— **Q** —— (e.g. (41a))
 c. —— ¬ Q —— (X) —— **Q** —— (e.g. (41b))
 d. —— **Q** —— (e.g. (40a))

In the sense that the achievement event (X) is present in some form in the temporal structures of (42a–c), the time of (X) might be seen in all of these to constitute the "time of situation" (TSit) in the framework of Klein (1994), with the difference that a gradual shift has occurred in "topic time" (TT) away from "time of situation". However, in (42d) there is no (X) with which the "time of situation" can even be identified, leaving no alternative but to assign the time of situation to the stative situation itself. Such is also the case for the examples in (39) having the structure

in (42c), where (X) may have actually occurred, but is excluded from the semantics of the verbal construction.

V-*te-iru* constructions with the meaning structure of (42d), and sometimes (42c), are therefore effectively decoupled from the lexical meaning of V: V-*te-iru* here no longer functions as a viewpoint aspect marker, but rather itself encodes the lexical aspectual meaning of the verb construction. These V-*te-iru* forms fill a semantic gap in the lexicon between verbs, strongly oriented toward non-stative, eventive meaning involving changeable situations, with few truly stative predicates, on the one hand, and adjective and copula forms, oriented toward more permanent stative meanings, such as properties of things, on the other. V-*te-iru* forms expressing the meaning in (42c) and (42d), by contrast, express states that are temporary and subject to change, but not highlighting change to the extent of garden-variety verbal event predicates. Note for example the more permanent character of the situation expressed by the copular construction in (43a) versus the more temporary character of the situation expressed by the corresponding V-*te-iru* construction in (43b).

(43) a. *Kono   heya   wa   kirei   da.*
     this   room   TOP   pretty   COP
     'This room is pretty/clean.'

   b. *Heya   wa   kirei   ni   nat-te-iru.*
     room   TOP   pretty   DAT   become-TE-IRU
     'The room is in a pretty/clean state.'

From a diachronic perspective, these V-*te-iru* forms represent the grammaticalization of temporary stative meaning in a category with a distinct formal character – that of the V-*te-iru* form – preserving certain verbal characteristics such as the *-ru/-ta* tense opposition, but having a semantic character with a closer affinity to adjectives and copular constructions.

7 Conclusion and future research perspectives

While the range of meanings encoded by viewpoint (grammatical) aspect in Japanese is characterized by a high degree of idiosyncrasy, compounded by a tendency for aspectual meaning to be combined with non-aspectual meaning, the meaning categories that comprise situation (lexical) aspect in Japanese exhibit a high degree of consistency with the fundamental lexical aspectual categories observed in English, and indeed all natural languages. At the most basic level, situation aspect is characterized by an opposition between homogeneous meaning (states and activities) and non-homogeneous meaning (achievements, semelfactives, and accomplishments), a

distinction that has been articulated in the Japanese grammatical tradition both in quantitative terms, highlighting the instantaneity vs. duration of a situation, as in the Kindaichi approach, or in qualitative terms, highlighting the presence vs. absence in aspectual structure of a component of change, as in the Okuda approach. Considered from a cognitive perspective, these two approaches are not mutually exclusive, but complementary to one another, as change in state is typically seen in human experience as forming an instantaneous boundary between two states, and lack of change by its very nature entails persistence over time of a state or activity.

While the fundamental categories of situation aspect may be universal in character, however, there are significant areas of divergence between languages in the membership of these categories, as is evident by comparing lexically corresponding forms in English and Japanese. Japanese, in contrast to English, is characterized by a poverty of truly stative verbs and a large number of change-of-state achievement verbs. Stative meaning in Japanese is primarily concentrated in adjective and copular forms, but we have proposed that alongside these are a group of V-*te-iru* forms that depart from the basic use of V-*te-iru* as a viewpoint aspect marker and lexically encode a variety of temporary stative meaning intermediate between permanent states, encoded by adjectives and copular forms, and changeable events, encoded by verbs. Diachronically, this use of V-*te-iru* may be seen as the grammaticalization of a morphosyntactic category distinct from the three preexisting categories of verbs, adjectives, and copular forms, but one that exhibits the general consistency seen in Japanese between morpho-syntactic form and lexical categories of aspectual meaning.

The temporal character of predicates interacts in fundamental ways with a wide variety of syntactic constructions that may not appear in themselves to express temporal categories of meaning, such as the transitive-intransitive alternations discussed in this chapter. A number of such interactions were noted by Kindaichi (1950) himself, who considers the relevance of his four-way classification to constructions such as V-*kakeru* 'be about to V, almost V' and V-*tai* 'want to V', but closer examination shows that semantic parameters such as intentionality (volitionality) are equally central to the behavior of constructions such as these as Kindaichi's aspectual categories. Whether and how parameters such as intentionality themselves interface with temporal categories is itself a topic ripe for future research (see Jacobsen 1997). It has at any rate become clear since the work of Tenny (1994) that aspectual meaning, in particular the notion of delimitation discussed in this chapter (whether or not an action or event culminates in an endpoint defining a change of state), is critical in defining the argument structure of predicates, and thereby governing the behavior of a variety of construction types in English and other languages. Following in the spirit of Dowty (1979), attempts have been made by Kageyama (1996) and others to formalize the Kindaichi/Vendler aspectual classes as conceptual schemas reflecting the temporal role of events and states in a way that captures their linkage to argument realization patterns in syntax. Kageyama,

for example, correlates the possibility of resultative patterns with the presence of resulting states in these conceptual schemas, as seen in the possibility of such patterns with achievements (e.g. *Garasu ga konagona ni ware-ta* [glass NOM pieces DAT smash$_{in}$-PST] 'The glass smashed to pieces') and accomplishments (e.g. *Garasu o konagona ni wat-ta* [glass ACC pieces DAT smash$_{tr}$-PST] '(He) smashed the glass to pieces'), but not with activities (e.g. **Kinzoku o taira ni tatai-ta* [metal ACC flat DAT pound-PST] '(He) pounded the metal flat' or **Syoozyo wa kutakuta ni nai-ta* [girl TOP exhaustion DAT cry-PST] 'The girl cried herself to exhaustion'). Aspectual categories thus play a central role in the way that arguments are realized or "linked" to the semantic structure of predicates (Levin and Rappaport Hovav 2005; Randall 2009). The degree to which such interfacing between temporal meaning, argument structure, and syntactic structure can be exploited to account for the syntactic behavior of various construction types in Japanese, English, and other languages is certain to constitute a rich field of aspect-related research in the years to come.

Acknowledgements

References

Bohnemeyer, Jürgen, and Mary Swift. 2004. Event realization and default aspect. *Linguistics and Philosophy* 27. 263–296.

Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge: Cambridge University Press.

Dowty, David. 1979. *Word meaning and Montague grammar.* Dordrecht: D. Reidel.

Jacobsen, Wesley M. 1984. Lexical aspect in Japanese. In David Testen, Veena Mishra and Joseph Drogo (eds.), *Papers from the Parasession on Lexical Semantics*, 150–161. Chicago: Chicago Linguistic Society.

Jacobsen, Wesley M. 1992. *The transitive structure of events in Japanese.* Tokyo: Kurosio Publishers.

Jacobsen, Wesley M. 1997. Agentivity and aspect in Japanese: A functional perspective. In Akio Kamio (ed.), *Directions in functional linguistics*, 83–116. Amsterdam: John Benjamins.

Jacobsen, Wesley M. 2005. On the fuzzy boundary between tense and aspect in Japanese. In Salikoko S. Mufwene, Elaine J. Francis and Rebecca S. Wheeler (eds.), *Polymorphous linguistics: Jim McCawley's legacy*, 261–282. Cambridge: MIT Press.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kindaichi, Haruhiko. 1950. Kokugo dōshi no ichibunrui [A classification of Japanese verbs]. *Gengo Kenkyu* 15. 41–63. Reprinted in Haruhiko Kindaichi (ed.), 1976. *Nihongo dōshi no asupekuto* [Aspect in Japanese Verbs], 5–26. Tokyo: Mugi Shobō.

Klein, Wolfgang. 1994. *Time in language*. New York: Routledge.

Levin, Beth and Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.

McClure, William. 1995. *Syntactic projections of the semantics of aspect*. Tokyo: Hituzi Syobo.

Nakatani, Kentaro. 2004. *Predicate concatenation: A study of the V te V predicate in Japanese*. Cambridge, MA: Harvard University dissertation.

Nakatani, Kentaro. this volume. Chapter 11: Complex predicates with *te* gerundive verbs. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Ogihara, Toshiyuki. 1999. Tense and aspect, In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 326–348. Oxford: Blackwell.

Okuda, Yasuo. 1978. Asupekuto no kenkyū o megutte [On the study of aspect] I, II. *Kokugo Kyōiku* 53. 33–44 / 54. 14–27.

Randall, Janet H. 2009. *Linking: The geometry of argument structure*. Dordrecht: Springer.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: Macmillan.

Searle, John R. 1983. *Intentionality*. Cambridge: Cambridge University Press.

Shirai, Yasuhiro. 2000. The semantics of the Japanese imperfective -*teiru*: An integrative approach. *Journal of Pragmatics* 32. 327–361.

Smith, Carlota. 1997. *The parameter of aspect*. Dordrecht: Kluwer Academic.

Soga, Matsuo. 1983. *Tense and aspect in modern colloquial Japanese*. Vancouver: University of British Columbia Press.

Takezawa, Koichi. this volume. Chapter 13: Inflection. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Taylor, Barry. 1977. Tense and Continuity. *Linguistics and Philosophy* 1. 199–220.

Tenny, Carol. 1994. *Aspectual roles and the syntax semantics interface*. Dordrecht: Kluwer Academic.

Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics*. Cambridge, MA: Blackwell.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160. Reprinted in Zeno Vendler (ed.), *Linguistics in philosophy* (1967). 97–121. Ithaca: Cornell University Press.

Williams, Edwin. 1981. Argument structure and morphology. *The Linguistic Review* 1. 81–114.

Hideki Kishimoto
16 Stative and existential/possessive predicates

1 Introduction

It is well observed that stative predicates have a unique status in the aspectual classi-fication of predicates in Japanese as well as in many other languages (Chapter 15 [Jacobsen, this volume]). This chapter explores the semantic and syntactic properties of the Japanese stative predicates that select two arguments, which are represented in many cases by a combination of dative case (experiencer, possessor) and nomina-tive case (theme), as in (1).

(1) a. *Ken ni wa sono kimoti ga   wakar-u.*
         Ken DAT TOP that feeling NOM understand-PRS
         'Ken can understand that feeling'

     b. *Ken ni wa mago   ga   i-ru.*
         Ken DAT TOP grandson NOM be.ANM PRS
         'Ken has a grandson'.

The fundamental question raised by such constructions is which of the two arguments serves as subject. Although most of the arguments that follow are syntactically oriented, the primary emphasis of this chapter is placed on how neatly the syntac-tic behavior of the predicates in question is correlated with their lexical-semantic properties.

    Japanese falls into the nominative-accusative language type, where the subject, but not the object, of a transitive verb is treated in the same way as the subject of an intransitive verb in case-marking terms. Nevertheless, it is also true that some predicates show idiosyncratic behavior not expected from this typological generaliza-tion. The varying behavioral patterns of predicates are closely tied to their semantic class, i.e. they are determined by the meanings they carry. In particular, two-place predicates encoding stative meanings mark their arguments differently from non-stative predicates, in such a way that the grammatical relations of the arguments cannot be uniquely identified by the surface forms. This gives rise to perennial issues of transitivity in Japanese generative grammar as well as in traditional Japanese grammar (called *Kokugogaku*).

    In terms of lexical categories, Japanese stative predicates can be divided into two broad classes. One is the verbal class (e.g. *dekiru* 'can do', *tigau* 'differ'), and the other the adjectival class comprising both adjectives (e.g. *kawaii* 'pretty) and

adjectival nouns (e.g. *hituyoo da* 'necessary').[1] Semantically, stative predicates are distinguished from non-stative (or eventive) predicates in their temporal properties. While non-stative predicates appearing in the present (or non-past) form describe events to take place in the future (or refer to habitual events), stative predicates in the present form describe present states. The two-place stative predicates may be classified into several types according to their case frames, as shown in Table 1.

Table 1: Semantic types of stative predicates

Predicate type	Basic case patterns	Examples
(I) Experiencer-subject predicates	<DAT, NOM>	*hituyoo da* 'necessary'
(II) Potential predicates	<DAT, NOM>, <NOM, ACC>	*wakaru* 'understand' *kakeru* 'can write'
(III) Predicates of liking/wanting	<NOM, NOM>	*suki da* 'be fond of', *hosii* 'want'
(IV) Existential predicates	<LOC, NOM>	*aru/iru* 'be'
(V) Possessive predicates	<DAT, NOM>	*aru/iru* 'have'

Notes: NOM = nominative, ACC = accusative, DAT = dative, LOC = locative

These dyadic stative predicates are peculiar in having "non-canonical" case patterns that deviate from the "canonical" case marking of <NOM (*ga*), ACC (*o*)> that is typically associated with dynamic (i.e. non-stative) transitive verbs. As shown in Table 1, several patterns of non-canonical case frames are distinguished, depending on the semantic properties of the predicates involved. While all these predicate types require two arguments – an experiencer and a theme, or a locative/possessor and a theme – they diverge in meaning and transitivity. The present chapter deals with the transitivity issues surrounding the dyadic stative predicates.

The transitivity of dyadic stative predicates is diagnosed by a number of syntactic tests identifying subject and/or object. The dyadic stative predicates in the categories of (I), (II) and (III) in Table 1 take the <DAT (*ni*), NOM (*ga*) or <NOM (*ga*), NOM (*ga*)> case-marking patterns (and some predicates can also take <NOM (*ga*), ACC (*o*)> case-marking patterns). It will be shown that these dyadic stative predicates function as "transitive" predicates, where the initial DAT argument or the initial NOM argument representing an experiencer is identified as subject and the second NOM argument

---

1 Adjectival nouns (*keiyōdōshi*) have a *na* ending in the adnominal form and a *da* ending in the con clusive form (e.g. *hituyoo na* and *hituyoo da*). Whether the *da* that is combined with adjectival noun stems should be regarded as an auxiliary (i.e. copula) or an inflectional suffix is an issue. Without repeating the arguments for each analysis, this chapter simply treasts the construction as a com bination of an adjectival noun stem followed by the copula. Nothing in the discussion of stative predicates here depends on the choice of analysis, however. For discussion of this issue, see Chapter 2 (Kishimoto and Uehara, this volume).

representing a theme is identified as object. Apart from the experiencer-subject predicates, existential (IV) and possessive (V) predicates take a non-canonical pattern of <DAT/LOC (*ni*), NOM (*ga*)>. Existential predicates are intransitive; the NOM argument serves as subject, while the LOC (*ni*) argument has the status of a locative adjunct syntactically. Possessive predicates are construed as "transitive", with the DAT (*ni*) argument identified as subject and the NOM argument as object. The predicates of possession count as transitive, regardless of whether they fall into the stative or non-stative class, and are considered to be derived via transitivization by reanalyzing the locative adjunct of an existential predicate as a possessor (the subject).

Since Kuno (1973), it has been commonly assumed that the non-canonical case-marking patterns are a hallmark of the stativity of two-place predicates. In actuality, however, the syntactic process of transitivization makes such case-marking patterns available even for non-stative predicates. Notably, the nominative arguments of possessive predicates, which count as objects by virtue of transitivization, are constrained by the so-called "definiteness restriction", which was motivated for the existential sentences in English (Milsark 1974, 1977). Nevertheless, on the "list" interpretation (cf. Rando and Napoli 1978), the same nominative arguments can accommodate definite expressions, apparently voiding the definiteness restriction.

The discussion in this chapter proceeds as follows. Section 2 illustrates how the case-marking patterns of stative predicates are regulated by their lexical meanings and how the grammatical roles of their arguments can be identified. Section 3 dicusses the transitivity issues of existential and possessive predicates. Section 4 shows that possessive predicates are derived from existential predicates via transitivization on the basis of the empirical observation that the nominative argument representing an object is constrained by the definiteness restriction. In Section 5, the list interpretation of possessive sentences is discussed. The present chapter concludes with perspectives for future research in Section 6. Overall, it will be shown that a semantically defined class of predicates exhibits more or less the same behavior in case marking and has many shared syntactic properties.

2 Transitivity of dyadic stative predicates

This section will first provide an overview of how the case marking of transitive stative predicates in Japanese is conditioned (Section 2.1), and then introduce a prominent issue surrounding the transitivity of dyadic stative predicates (Section 2.2). Since the two arguments of dyadic stative predicates carry non-canonical case markings that diverge from the canonical nominative-accusative marking, the issue arises as to whether they count as transitive or intransitive: if they take both subject and object, they are transitive, but if they select two subjects (with no object), they

are intransitive. A number of subject and object tests are provided to show that dyadic stative predicates are indeed transitive.

2.1 Basic case frames of dyadic stative predicates

While the regular case-marking pattern for ordinary transitive predicates is <NOM (*ga*), ACC (*o*)>, as in (2a), the dyadic (two-place) stative predicates treated in this chapter take different case-marking patterns of <DAT (*ni*), NOM (*ga*)> or <NOM (*ga*), NOM (*ga*)>, as exemplified by (2b).

(2) a. *Ken*   ***ga***   *sono*   *hon*   ***o***   *yon-da*.
     Ken   NOM   that   book   ACC   read-PST
     'Ken read that book.'

   b. *Ken*   {***ni***/***ga***}   *sono*   *hon*   ***ga***   *hituyoo*   *da*.
     Ken   {DAT/NOM}   that   book   NOM   necessary   COP
     'Ken needs that book.'

In (2b), the initial argument *Ken* denoting an experiencer may be marked with either dative or nominative case, while the second argument *sono hon* 'that book' denoting a theme is invariably marked in the nominative.[2] It is shown below that the initial argument functions as the subject of the sentence regardless of its case marking, and the second argument as object. In other words, sentences like (2b) are "transitive", although their transitivity is not so high as the ordinary transitive predicates that take the <NOM, ACC> case pattern. By virtue of its low transitivity, sentence (2b) cannot mark its theme ('that book') in the accusative, even though it is identified as an object, as the ungrammaticality of (3) shows.

(3) **Ken*   *ga*   *sono*   *hon*   *o*   *hituyoo*   *da*.
     Ken   NOM   that   book   ACC   necessary   COP
     'Ken needs that book.'

     The stative transitive predicates taking the non-canonical case-marking patterns share the semantic property of representing states rather than dynamic events. The states denoted by them cover such static notions as possession, perception, and necessity, shown as 'x is in the state of possessing/perceiving/needing y', and their

---

2 Some stative predicates like *suki da* 'be fond of' do not allow the dative case marking on the experiencer argument, and it seems that what type of predicate excludes a dative marked subject is often subject to idiolectal variation, as noted by Shibatani (1978).

two arguments, *x* and *y*, are aligned with their respective thematic roles and case marking as follows.

(4) Lexical meaning: 'x is in the state of possessing/perceiving/needing y'

| |

  Thematic role: Possessor/Experiencer   Theme

| |

  Case marking: NOM/DAT   NOM

According to Shibatani (2001), the predicates displaying <DAT/NOM, NOM> or <NOM, NOM> case-marking patterns can be classified according to their meanings, as in Table 2.

Table 2: Semantic classification of two-place stative predicates

(A) possession and existence	*ooi* 'many', *sukunai* 'few', *syoosuu da* 'a few', *aru/iru* 'have, be'
(B) psychological states	*suki da* 'fond', *kowai* 'fearful', *kirai da* 'disliking', *kyoohu da* 'fear', *kutuu da* 'painful'
(C) physiological perception[3]	*itai* 'painful', *kayui* 'ichty', *(onaka ga) suku* 'hungry', *zuki zuki suru* 'hurt'', *hiri hiri suru* 'sting'
(D) visual or auditory perception	*mieru* 'visible', *kikoeru* 'audible'
(E) necessity and desire	*iru* 'need', *hituyoo da* 'necessary', *hosii* 'want', *nomi tai* 'want to drink'
(F) potentiality and ability	*yomeru* 'readable, legible', *wakaru* 'understandable', *rikai dekiru* 'can understand'

In terms of lexical categories, these predicates include verbs, adjectives, and adjectival nouns, and in terms of lexical strata, they include native Japanese, Sino-Japanese, and mimetic words. The number of simple stative verbs is fairly small,

---

3 Predicates expressing physiological states only take <NOM, NOM> patterns, and the first argument cannot be marked with genitive case.

(i) a. *Ken ga   onaka   ga   ita i.*

     Ken NOM stomach NOM painful PRS

     'Ken has a pain in his stomach.'

   b. **Ken no   onaka   ga   ita i.*

     Ken GEN stomach NOM painful PRS

     'Ken's stomach has a pain.'

This fact suggests that a physiological predicate like *itai* 'painful' is a two place predicate. The tran sitivity of the physiological predicates cannot be determined with recourse to the subject and object diagnostics to be discussed in this chapter, however. The reason is that physiological predicates select body part expressions as their second arguments. The subject and object diagnostics target human nouns, and thus cannot be applied to the second arguments of physiological predicates.

but the actual inventory of stative verbs taking the non-canonical case patterns is quite large, because complex potential verbs belonging to the stative class can be productively formed with the potential affix *−(r)are* or *−(r)e*.

Several exceptions to the general rule on the case marking of stative predicates are found. While dyadic stative predicates generally do not allow the <NOM, ACC> case pattern, stative predicates expressing the meaning of liking/disliking and wanting, including *suki da* 'be fond of', *kirai da* 'dislike' and *hosii* 'want', do allow the <NOM, ACC> case pattern, alongside the <NOM, NOM> case pattern, as exemplified in (5).

(5) *Ken   ga   Mari   {ga/o}   suki   da.*
     Ken   NOM   Mari   {NOM/ACC}   like   COP
     'Ken likes Mari.'

Although semantic, pragmatic, or stylistic factors appear to govern the choice or preference of case marking for the theme argument (see Shibatani 1978; Iori 1995), the logical meaning of the sentence is essentially the same, regardless of the choice of case marking on the theme. Accordingly, the following relation of argument realization can be postulated for this class of stative predicates.

(6) Lexical meaning: 'x is in the state of liking/wanting y'
                          | |
     Thematic role: Experiencer   Theme
                          | |
     Case marking: NOM   NOM/ACC

Two-place potential verbs derived via suffixation of the potential affixes *−(r)e/-(r)are* are allowed to mark the second theme argument with accusative case when the first experiencer argument is marked with nominative case, but not when it is marked with dative case, as exemplified in (7).

(7) a. *Ken   ga   sono   uta   {o/ga}   uta-e-ru.*
         Ken   NOM   that   song   {ACC/NOM}   sing-POTEN PRS
         'Ken can sing that song.'

     b. *Ken   ni   sono   uta   {ga/*o}   uta-e-ru.*
         Ken   DAT   that   song   {NOM/ACC}   sing-POTEN PRS
         'Ken can do that song.'

With potential verbs, the accusative marking of the second theme argument is possible to the extent that the volitional control meaning of the original transitive verbs that take the <NOM, ACC> case frame is preserved. (8) represents a mapping relation between the lexical meaning and case marking of this class of predicates.

(8) Lexical meaning: 'x is capable of doing y'
                        | |
     Thematic role: Experiencer   Theme
                        | |
     Case marking: NOM/DAT   NOM/ACC

Japanese has the case-marking constraint requiring that there be at least one nominative argument per tensed clause (Shibatani 1978).[4] With two-place potential predicates, the initial experiencer argument can be marked with either dative or nominative case, and the second theme argument with nominative or accusative case. Logically speaking, two-place potential predicates should be allowed to take the case frame <DAT, ACC>, alongside the <DAT/NOM, NOM> and <NOM, ACC> frames, but the nominative-case constraint prohibits the <DAT, ACC> case array, as shown by the unavailability of the "*X ni Y o*" pattern in (7b).

It is worth noting that predicates like *kikoeru* 'hear' and *mieru* 'see', which are spontaneous forms of *kiku* 'listen to' and *miru* 'look at', allow only the <DAT/NOM, NOM> patterns, as shown in (9).

(9) a. *Ken {ni/ga}   (*reiseini) sono ongaku ga   kikoe-ru.*
         Ken {DAT/NOM} (calmly)   that music   NOM   hear-PRS
         'Ken hears that music (*calmly).'

     b. **Ken ga   sono ongaku o   kikoe-ru.*
         Ken NOM   that music   ACC hear-PRS
         'Ken hears that music.'

The <NOM, ACC> case frame is not allowed for *kikoeru*, as in (9b), because the spontaneous verb does not carry the meaning of subject's volitional control. The absence of volitional control meaning is confirmed by the fact that *kikoeru* 'hear' in (9a) is not compatible with an adverb like *reiseini* 'calmly', a modifier that describes a manner in which a (controlled) act is done. In contradistinction, the perception verbs *kikeru* 'can listen to' and *mirareru* 'can look at', derived from *kiku* 'listen to' and *miru* 'look at' through the addition of the potential suffixes *–(r)are/-(r)e*, display full-fledged case-marking patterns.[5]

---

4 It has been observed that the nominative case constraint does not apply to certain types of clauses (Shibatani 1978; Kuroda 1978; Kishimoto 2010). See Kishimoto (forthcoming) for a proposal to account for why some clauses are constrained by the nominative case constraint, but others are not.
5 The perception verb *mirareru* has a short form *mireru*. Although *mireru* is sometimes claimed to be a non standard form, it is very often used in colloquial Japanese.

(10) a. *Ken   {ni/ga}   (reiseini)   sono   ongaku   ga   kik-e-ru.*
      Ken   {DAT/NOM}   (calmly)   that   music   NOM   listen-POTEN PRS
      'Ken can listen to that music (calmly).'

    b. *Ken   ga   (reiseini)   sono   ongaku   o   kik-e-ru.*
      Ken   NOM   (calmly)   that   music   ACC   listen-POTEN PRS
      'Ken can listen to that music (calmly).'

The availability of the case-marking pattern <NOM, ACC> in (10b) is due to the fact that the potential form of the perception verb carries the meaning of volitional control that the base verb has. The presence of the control meaning can easily be discerned by the fact that the clause with *kikeru* 'can listen to' can accommodate the adverb *reiseini* 'calmly' regardless of its case-marking pattern, as in (10).

One-place potential predicates, unlike two-place potential predicates, do not display varying case-marking patterns. In the case of one-place potential verbs (with no theme or path arguments), the subject must appear in the nominative case, and not in the dative case.

(11) *Kono   hito   {ga/*ni}   hatarak-e-ru.*
    this   person   {NOM/DAT}   work-POTEN PRS
    'This person can work.'

The dative marking on the sole argument in (11) is excluded by the nominative-case constraint requiring at least one nominative case per sentence.[6] The mapping relation of an intransitive potential predicate like *hatarakeru* 'can work' is represented in (12).

---

6 Despite this generalization, a dative subject clause without any nominative argument is fairly acceptable on a contrastive interpretation.

(i) *(Kare   wa   oyog e na i   ga)   ano   hito   ni   wa   oyog e ru.*
   he   TOP   swim POTEN-NEG-PRS   but   that   man   DAT   TOP   swim POTEN-PRS
   '(He cannot swim, but) that man can swim.'

Dative marking can often be replaced by a complex postposition like *nitotte* 'for', perhaps for the purpose of adding an emphasis, as in (ii).

(ii) *Kare   {ni/nitotte}   wa   sore   ga   hituyoo   da.*
   he   DAT/for   TOP   that   NOM   necessary   COP
   'That is necessary for him.'

In the light of this fact, I surmise that *ni* appearing on the dative subject could be a postposition, given a contrastive context, in which case the nominative case constraint is not in force (cf. Kishimoto forthcoming).

(12) Lexical Meaning: 'x is capable of doing work'
                                     |
      Thematic role: Experiencer
                                     |
      Case marking: NOM (*DAT)

Note that the case-marking patterns available for transitive potential predicates – i.e. <NOM, NOM/ACC> and <DAT, NOM> – are possible with intransitive verbs suffixed with the potential affix if they take a traverse path argument, which is marked with accusative case in their non-potential counterparts.

(13) a. *Ken   ga   sono   miti   {o/ga}   aruk-e-ru.*
          Ken   NOM   that   path   {ACC/NOM}   walk-POTEN PRS
          'Ken can walk that path.'

      b. *Ken   ni   sono   miti   {ga/*o}   aruk-e-ru.*
          Ken   DAT   that   path   {NOM/ACC}   walk-POTEN PRS
          'Ken can walk that path.'

Apparently, intransitive verbs are not turned into transitive predicates even if they take accusative path arguments (cf. Haig 1981). Nevertheless, when intransitive verbs taking traverse path arguments are formed into potential predicates, they behave like transitive potential predicates in case-marking terms.

   Finally, predicates like *sinsetu da* 'be kind', *yasasii* 'be kind', *isogasii* 'be busy', *mero-mero da* 'have a crush on' which describe the subject's attitude (or mental state) take a <NOM, DAT> case pattern (Masuoka 1987).

(14) *Ken$_i$ wa   zibun$_i$ no   kodomo   dake   ni   yasasi-i.*
      Ken   TOP   self   GEN   child   only   DAT   kind-PRS
      'Ken is kind only to his children.'

The lexical meaning of the adjective *yasasii* 'be kind', represented as 'x is kind to y', has the following mapping relation to the thematic roles and case marking.

(15) Lexical meaning: 'x   is kind to   y'
                                  | |
      Thematic role: Experiencer   Theme
                                  | |
      Case marking: NOM   DAT

The dative argument of *yasasii* 'be kind' counts as a complement specifying the target toward which the subject's kindness is directed. With this class of predicates,

the nominative argument serves as a subject, and the dative argument a complement, as confirmed by the fact that subject-oriented *zibun* takes the nominative experiencer as its antecedent in (14).

To summarize, dyadic stative predicates are grouped into several types that are largely determined by the kinds of semantic roles their arguments bear. The dyadic stative predicates show varying case-marking patterns, as summarized in Table 3.

Table 3: Possible case-marking patterns of dyadic stative predicates

Predicates	Case-Marking Patterns
hituyoo da 'necessary'	<DAT, NOM>, <NOM, NOM>
suki da 'like'	<NOM, NOM>, <NOM, ACC>
yom e ru 'can read'	<DAT, NOM>, <NOM, NOM>, <NOM, ACC>
sinsetu da 'kind'	<NOM, DAT>

In the next section, I will turn to discussion of the long-standing "transitivity" issues surrounding the status of the stative predicates taking non-canonical case-marking patterns, with particular emphasis on the question of how the grammatical relations of the arguments of dyadic stative predicates can be empirically determined.

2.2 Experiencer arguments as the syntactic subject

While non-stative two-place predicates take a <NOM, ACC> case marking pattern, potential predicates taking two arguments of experiencer and theme allow three distinct case-marking patterns of <NOM, ACC>, <DAT, NOM>, and <NOM, NOM>. Nominative case is often considered to be the default case marking assigned to subjects, and subjects are not marked with accusative case (in matrix clauses).[7] Accordingly, two-place stative predicates with the <NOM, ACC> case-marking pattern are considered to be transitive, taking both subject and object. On the other hand, the non-canonical case-marking patterns of two-place stative predicates which have

---

7 In Japanese, accusative case may be assigned to the subject of the embedded predicate in the exceptional case marking (ECM) and small clause (SC) constructions, as in (i).

(i) a. *Ken wa [Mari o   kawai i   to] omot ta.*
　　　Ken TOP Mari ACC cute PRS that think PST
　　　'Ken thought Mari to be cute.' (ECM construction)

　　b. *Ken wa [Mari o   kawaiku] omot te   i ru.*
　　　Ken TOP Mari ACC pretty   think GER be PRS
　　　'Ken thinks Mari pretty.' (SC construction)

In both constructions, the matrix predicate can be assumed to license the accusative case marking on the embedded subject.

the second theme argument marked with nominative case, i.e. <DAT/NOM, NOM>, have led to a debate over whether these predicates should be construed as intransitive predicates with no objects or transitive predicates with objects.

Most researchers in traditional Japanese grammar view nominative-marked nominals uniformly as subjects; under this view, the second nominative argument of a dyadic stative predicate would be regarded as representing a subject (Hashimoto 1969; Onoe 1997–1998; and others). There is also a view taking the second nominative argument to assume a third type of grammatical relation different from subject or object (Tokieda 1950).[8] In these analyses, the second nominative argument of stative predicates cannot be an "object". Studies in Japanese generative grammar in the 1970s and 1980s make a marked contrast in this respect. Kuno (1973) and Shibatani (1978) claim that the nominative case marking may be used to represent an object, and that the second nominative argument of a dyadic stative predicate is an object.

To be more concrete, in the traditional approach taking the nominative case to uniformly signal the subject status, (16a) should be treated as an intransitive sentence with two nominative subjects (i.e. a double subject construction), and by extension, (16b) would presumably be analyzed in the same way, a position Shibatani (2001) has argued for.

(16) a. *Ken   ga   Mari   ga   suki   da.*
          Ken   NOM   Mari   NOM   fond   COP
          'Ken likes Mari.'

   b. *Ken   ni   Mari   ga   hituyoo   da.*
          Ken   DAT   Mari   NOM   necessary   COP
          'Ken needs Mari.'

In the early generative approach taking the second nominative argument of a two-place stative predicate as an object (Kuno 1973; Shibatani 1978), by contrast, the double nominative construction in (16a) and the dative-nominative construction in (16b) are identified as transitive sentences in which the first argument is subject and the second is object (see also Kuno and Johnson 2005).

The divergence of views is further compounded by the "major subject" construction, which often displays the same case-marking patterns superficially as those of

---

8 Apart from the issue on the status of subjects in stative predicates, many different ways of characterizing subjects have been proposed by traditional Japanese grammarians (e.g. Matsushita 1928; Yamada 1936; Kitahara 1981). Besides, some scholars (e.g. Watanabe 1971) do not commit themselves to assigning special status to subjects, and others (e.g. Mikami 1953) even deny the relevance of the notion of subject to Japanese grammar.

dyadic stative predicates. In the major subject construction, an extra nominative-marked argument, which is not selected by the predicate, shows up alongside the arguments selected by it. Observe (17a).

(17) a. *Ken   ga   kodomo   ga   (benkyoo   ni)   isogasi-i.*
        Ken   NOM   child   NOM   (study   DAT)   busy-PRS
        As for Ken, his child is busy (with his study).'

    b. *[Ken   no   kodomo]   ga   isogasi-i.*
        Ken   GEN   child   NOM   busy-PRS
        'Ken's child is busy.'

In (17a), the initial nominative argument is identified as a major subject, which can be realized alternatively as a genitive argument embedded inside the thematic subject, as in (17b). In contrast, the first nominative argument in (16a) cannot be realized as a genitive argument, as seen from the ungrammaticality of (18).

(18) **[Ken   no   Mari]   ga   suki   da.*
      Ken   GEN   Mari   NOM   fond   COP
      (lit.) 'Ken's Mari is fond.'

The availability of the alternative form in (17b) indicates that the first nominative argument in (17a) is not selected by the verb, but rather is licensed with an "aboutness" relation (specifically, a possession relation) with the second nominative argument, which is the thematic subject of the predicate (Kuno 1973).

    Given that the nominative case is capable of coding a number of different grammatical relations, it is instructive to ask whether there are any tests that allow us to assess the grammatical relations of the arguments of stative predicates. Japanese does indeed have a battery of diagnostic tests that can be used to assess the subject and/or object status of arguments. These diagnoses will help us identify the initial dative/nominative experiencer arguments in (16) as subjects and the second nominative theme arguments as objects.

    Let us begin with three diagnoses of subjecthood, namely reflexivization, controlled PRO, and PRO with arbitrary interpretation (PRO$_{arb}$) (Shibatani 1984; Kishimoto 2000, 2005). First, in an ordinary transitive clause like (19), the subject-oriented pronoun *zibun* 'self' can take the nominative subject, but not the accusative object, as its antecedent.

(19) *Ken$_i$ ga   Mari$_j$ o   zibun$_{i/*j}$ no   heya   de   home-ta.*
      Ken   NOM   Mari   ACC   self   GEN   room   in   praise-PST
      (lit.) 'Ken praised Mari in self's room.'

With a stative predicate taking the <DAT/NOM, NOM> case-marking pattern, the first dative/nominative argument representing an experiencer – but not the second nominative argument representing a theme – can serve as the antecedent of *zibun*, as (20) shows.

(20) *Ken$_i$ {ni/ga}   Mari$_j$ ga zibun$_{i/*j}$ no keikaku ni hituyoo   dat-ta.*
     Ken {DAT/NOM} Mari NOM self   GEN plan   for necessary COP PST
     (lit.) 'Ken needed Mari for self's plan.'

This fact suggests that the first experiencer argument, not the second theme argument, should serve as a subject.

     Second, controlled PRO can occur only in subject position. Thus, the nominative subject, but not the accusative object, of a regular transitive predicate can be rendered as PRO (via embedding under a desiderative predicate like *hosii* 'want'), as in (21).

(21) a. *Mari wa Ken$_i$ ni   [PRO$_i$ kodomo o   home-te]   hosikat-ta.*
        Mari TOP Ken DAT   child ACC praise-GER want-PST
        'Mari wanted Ken to praise his children.'

     b. **Mari wa kodomo$_i$ ni   [Ken ga   PRO$_i$ home-te]   hosikat-ta.*
        Mari TOP child   DAT Ken NOM   praise-GER want-PST
        (lit.) 'Mari wanted the children$_i$ for Ken to praise PRO$_i$.'

As seen in (21), the dative phrase in the matrix clause can control PRO, which can appear only in the embedded subject position. With dyadic stative predicates like *suki da* 'be fond of', only the first experiencer argument, and not the second theme argument, can be replaced by controlled PRO, as seen in (22).

(22) a. *Mari wa Ken$_i$ ni   [PRO$_i$ kodomo ga   suki de   at-te] hosikat-ta.*
        Mari TOP Ken DAT   child NOM fond COP be-GER want-PST
        'Mari wanted Ken to like children.'

     b. **Mari wa kodomo$_i$ ni   [Ken ga   PRO$_i$ suki de   at-te] hosikat-ta.*
        Mari TOP child   DAT Ken NOM   fond COP be-GER want-PST
        (lit.) 'Mari wanted the children$_i$ for Ken to like PRO$_i$.'

With the predicate *suki da* 'fond', the first experiencer argument can be tuned into PRO controlled by the dative argument to *hosii*, and thus, (22a) is a possible structure. By contrast, in (22b), the theme argument is controlled by the dative argument, so (22b) is not acceptable.[9] Since PRO can appear only in subject position, this fact

---

9 Even when a dative subject is embedded, its status can still be confirmed, since the second argument cannot be replaced by PRO controlled by the dative (controller) argument of the matrix clause.

shows that the first experiencer argument, but not the second theme argument, serves as a subject.

Third, PRO_{arb}, just like controlled PRO, can only stand in subject position. As illustrated in (23), PRO_{arb}, can show up in the nominative subject position, but not in the accusative object position, of a regular transitive verb.

(23) a. *[PRO_{arb} kodomo o homeru] koto wa ii koto da.*
      child ACC praise that TOP good thing COP
     'It is a good thing PRO_{arb} to praise children.'

   b. **[Ken ga PRO_{arb} homeru] koto wa ii koto da.*
     Ken NOM   praise thing TOP good thing COP
     'It is a good thing for Ken to praise PRO_{arb}.'

With a dyadic stative predicate like *suki da* 'be fond of', the first experiencer argument, but not the second theme argument, can be replaced by PRO_{arb}, as (24) shows.

(24) a. *[PRO_{arb} kodomo ga suki na] koto wa ii koto da.*
      children NOM fond COP thing TOP good thing COP
     'It is a good thing PRO_{arb} to like children.'

   b. **[Ken ga PRO_{arb} suki na] koto wa ii koto da.*
     Ken NOM   fond COP thing TOP good thing COP
     'It is a good thing for Ken to like PRO_{arb}.'

This fact suggests, again, that the first experiencer argument of a two-place stative predicate should count as a subject. Since all the syntactic operations discussed above target subjects, the data lead to the conclusion that the first dative/nominative experiencer argument should be construed as subject in dyadic stative constructions.

Let us turn to the discussion of how objects are diagnosed. One object test is based on the insertion of the formal noun *koto*, which does not carry any substantive lexical meaning (see Sasaguri 2000; Kishimoto 2004; Takubo 2010).

(25) a. *Ken ga Mari (no koto) o damasi-ta.*
     Ken NOM Mari (GEN fact) ACC deceive-PST
     'Ken deceived Mari.'

   b. *Ken (*no koto) ga Mari o damasi-ta.*
     Ken (GEN fact) NOM Mari ACC deceive-PST
     'Ken deceived Mari.'

   c. *Mari (*no koto) ga Ken niyotte damas-are-ta.*
     Mari (GEN fact) NOM Ken by   deceive-PASS PST
     'Mari was deceived by Ken.'

In (25a), the object may be realized merely as *Mari o* [Mari ACC], where the accusative particle is attached to the noun *Mari*, or as a complex phrase *Mari no koto o* [Mari GEN fact ACC], where the "formal noun" *koto* is added after *Mari*, and the presence or absence of the formal noun does not affect the cognitive meaning of the sentence. The attachment of *koto*, which is strictly limited to the surface object position, is not feasible with the subject *Ken ga* [Ken NOM] in (25b). In fact, if the object *Mari* is promoted to subject by passivization, as in (25c), it is no longer compatible with with the formal noun. This grammatical restriction on the surface position of the formal noun *koto* allows us to determine which argument of a dyadic stative predicate like *wakaru* 'recognize' serves as object. Consider (26).

(26) a. *Ken   {ni/ga}   Mari (no   koto)   ga   wakat-ta.*
        Ken   {DAT/NOM}   Mari (GEN fact)   NOM   recognize-PST
        'Ken was able to recognized Mari.'

    b. *Ken   (*no   koto)   {ni/ga}   Mari   ga   wakat-ta.*
        Ken   (GEN   fact)   {DAT/NOM}   Mari   NOM   recognize-PST
        'Ken was able to recognize Mari.'

A comparison of the data in (26) suggests that the second nominative argument, but not the first dative/nominative argument, of *wakaru* qualifies as an object.

In this connection, note that the major-subject construction often comprises more than one nominative argument, as in (27), where the first nominative argument is a major subject licensed with an "aboutness" relation to the thematic subject of the verb.

(27) *Ken   ga   kodomo   ga   benkyoo   ni   isogasi-i.*
     Ken   NOM   child   NOM   study   DAT   busy-PRS
     'As for Ken, his child is busy studying.'

In (27), the second nominative argument *kodomo* 'child' can be the antecedent of the subject-oriented *zibun*, and does not tolerate the addition of the formal noun *koto*.

(28) a. *Ken   ga   kodomo$_i$ ga   zibun$_i$ no   benkyoo   ni   isogasi-i.*
        Ken   NOM   child   NOM   self   GEN   study   DAT   busy-PRS
        'As for Ken, his child$_i$ is busy (for his$_i$ study).'

    b. *Ken   ga   kodomo   (*no   koto)   ga   isogasi-i.*
        Ken   NOM   child   GEN   fact   NOM   busy-PRS
        'As for Ken, his child is busy.'

The examples in (28) show that the second nominative argument in the major subject construction in (27) counts as a subject rather than as an object. It should be

apparent then that the major-subject construction cannot be treated on a par with non-canonical case-marking constructions derived by dyadic stative predicates.

In short, a battery of subject and object tests shows that the first dative/nominative arguments of dyadic stative predicates are construed as subjects, and the second nominative arguments as objects. Dyadic stative predicates displaying the non-canonical case-marking patterns are transitive with the subject-object structure.

3 Existential and possessive predicates

Stative predicates allow a locative adjunct to be marked with *ni*. Since the locative marker is homonymous with the dative marker *ni*, certain intransitive stative predicates taking locative adjuncts display the same "*ni-ga*" case patterns as do transitive stative predicates – a phenomenon that raises another transitivity issue. This problem has often been discussed in the light of the existential/possessive verbs *aru* and *iru* because these predicates select *ni*-marked arguments, which may represent either locative adjuncts or dative subjects. In the following discussion, I will provide an illustration of how the transitivity of the existential/possessive verbs *aru* and *iru* is determined. (This chapter does not treat the dialectal verb *oru* 'be', which can fill some slots in the inflectional paradigms of the animate *iru* in Standard Japanese: see Kinsui 2006).

To begin, it is appropriate to account for the semantic properties of the two verbs. When *aru* and *iru* carry an existential meaning 'x is located at y', they specify the spatial relation of an entity *x* relative to a location *y*. On the possessive interpretation 'x possesses y', on the other hand, the same verbs express an abstract relation of possession that a human individual *x* has with respect to an entity *y*. Note that the location argument in the existential construction typically though not exclusively refers to an inanimate entity or place, and the possessor argument in the possessive construction typically though not exclusively refers to an animate entity or a human individual, although both arguments are marked with *ni*. (A possessor needs to be animate, unless a part-whole relation is expressed, in which case the possessor can be inanimate.)

Aru and *iru* are the only pair of verbs in Japanese that alternate according to the animacy of the nominative argument. These verbs display different alternation patterns, depending on whether they express an existential or possessive meaning. When used as existential predicates, the verbs must match the nominative argument, i.e. the nominative argument must be animate for *iru* and inanimate for *aru*, as in (29).

(29) a. *Kooen   ni   {kodomo/*benti}   ga   i-ru.*
          park   LOC   {child/bench}   NOM   be.ANM PRS
          'There were {children/benches} in the park.'

b. *Kooen   ni   {*kodomo/benti}   ga   ar-u.*
park   LOC   {child/bench}   NOM   be.INAN PRS
'There were {children/benches} in the park.'

When *aru* is used as a possessive verb, the nominative argument can be animate or inanimate, but the possessive *iru* permits only an animate noun as its nominative argument, as shown in (30).

(30) a. *Sensei   ni   {kodomo/*okane}   ga   i-ru.*
teacher   DAT   {child/money}   NOM   have.ANM PRS
'The teacher has {children/money}.'

b. *Sensei   ni   {kodomo/okane}   ga   ar-u.*
teacher   DAT   {child/money}   NOM   have.INAN PRS
'The teacher has {children/money}.'

Note incidentally that *okane* 'money' in (30a) would pose no problem on an irrelevant reading where *iru* is taken as a homonymous verb meaning 'need'.

The verbs *aru* and *iru* superficially show the same case-marking pattern of <DAT or LOC *ni*, NOM *ga*> irrespective of whether they express a possessive or locative meaning. They differ, however, in transitivity (Kuno 1973; Shibatani 1978; Kishimoto 2000, 2005, 2012). The transitivity of *aru* and *iru* is determined by the meaning they carry. On the one hand, when a locative/existential meaning is expressed, they function as intransitive predicates, with the *ni*-marked locative serving as an adjunct, as in (31).

(31) Lexical meaning: 'x   is located at   y'
                          | |
     Grammatical relation: SBJ   ADJUNCT
                          | |
     Case marking: NOM   LOC (*ni*)

Note that the basic word order of existential sentences is 'locative-nominative' (see Kuno 1973). On the possessive reading, on the other hand, *aru* and *iru* behave as transitive predicates, and the first *ni*-marked possessor argument functions as a subject.

(32) Lexical meaning: 'x   possesses   y'
                          | |
     Grammatical relation: SBJ   OBJ
                          | |
     Case marking: DAT   NOM

As exemplified by (29) and (30) above, the choice between *aru* and *iru* is determined by the animacy of the nominative argument, and this fact raises a further issue over how the alternation of the two verbs is determined. A classic solution is to treat *aru* and *iru* as a pair of alternating verbs when they are intransitive, while treating *aru* as the only transitive possessive verb (Kuno 1973; Shibatani 1978, 1990). Under this analysis, *iru* is invariably construed as an intransitive existential verb, regardless of its meaning (existence or possession), and the verb alternation is regarded as an instance of subject agreement. Thus, *Sensei ni kodomo ga iru* [teacher DAT child NOM be.ANM] 'The teacher has children.' (cf. (30)) is construed as an existential sentence, despite its possessive interpretation. Shibatani (1978) attempts to justify this analysis in light of the contrast in acceptability between (33a) and (33b).

(33) a. *Kimi   ni   wa   rippana   go-ryoosin   ga   irassyar-ru.*
     you   DAT TOP excellent HON-parents NOM be.ANM.HON PRS
     'You have excellent parents.'

    b. **Kimi   ni   wa   rippana   go-ryoosin   ga   o-ari-ni-nar-u.*
     you   DAT TOP excellent HON-parents NOM have.INAN.HON PRS
     'You have excellent parents.'

(33a) is acceptable because the honorification of the verb *iru* is targeted at the nominative argument ('parents'). (33b), on the other hand, is unacceptable because the honorific verb *aru* infelicitously agrees with the sentence-initial personal pronoun *kimi* 'you', which can only be used to refer to a person of equal or lower social status. In (33b), *kimi* cannot be targeted for subject honorification for the reason that the referent would have to be socially superior to the speaker (Harada 1976). In light of the contrast in (33), Shibatani claims that *iru* is an intransitive verb whose subject shows up to its immediate left and cannot be a transitive verb. A scrutiny of additional examples, however, suggests that the transitive use is available with *iru*, as shown by (34), where the honorification of the animate *iru* expressing a possessive meaning is anchored by the dative argument 'Professor Abe'.

(34) *Abe-sensei   ni   wa   annani kawaii akatyan ga   irassyar-u.*
    Abe-teacher DAT TOP such   cute   baby   NOM have.ANM.HON PRS
    'Professor Abe has such a cute baby.'

Since this dative argument is shown to be a subject on independent grounds, as discussed below, it is plausible to state that *iru* has a transitive use, despite Shibatani's claim to the contrary (cf. Tsunoda 2009).

    There is in fact a sense in which both *aru* and *iru* are used transitively when they express a possessive meaning, because, in such cases, the dative argument is diagnosed as subject by all the subject diagnoses discussed above (Kishimoto 2000). First, when *iru*, as well as *aru*, is used existentially, the nominative argument

is diagnosed as subject by the subject tests (i.e. reflexivization, controlled PRO, and PRO$_{arb}$), as seen from (35).

(35) a. *Zibun$_i$ no   heya   ni   Ken$_i$ ga   i-ru.*
          self   GEN   room   LOC   Ken   NOM   be.ANM PRS
          'Ken is in his own room.'

    b. *Sensei   wa   Ken$_i$ ni   [koko   ni   PRO$_i$ i-te]   hosikat-ta*
          teacher TOP   Ken   DAT   here   LOC   be.ANM GER   want-PST
          'The teacher wanted Ken to be here.

    c. *[Asoko   ni   PRO$_{arb}$ iru]   koto   wa   ii   koto   da.*
          there   LOC   be.ANM   that   TOP   good   thing   COP
          'It is a good thing PRO$_{arb}$ to be there.'

In contrast, the same diagnoses pick out the dative argument as subject when *aru* and *iru* encode a possessive meaning. Remarkably, the same results obtain regardless of the choice of verb, as shown in (36).

(36) a. *Ken$_i$ ni   zibun$_i$ no   kodomo   ga   {ar-u/i-ru}.*
          Ken   DAT   self   GEN   child   NOM   {have.INAN PRS/have.ANM PRS}
          'Ken also has his own child.'

    b. *Mari   wa   Ken$_i$ ni   [PRO$_i$ kodomo   ga*
          Mari   TOP   Ken   DAT   child   NOM

        {i-te/at-te}]   hosikat-ta.
          {have.ANM GER/have.INAN GER} want-PST
          'Mari wanted Ken to have children.'

    c. *[PRO$_{arb}$ kodomo   ga   {iru/aru}]   koto   wa   ii   koto   da.*
                  child   NOM   {have.ANM/have.INAN}   that   TOP   good   thing   COP
          'It is a good thing PRO$_{arb}$ to have children.'

Furthermore, the subject-honorific verb *irassyaru* (for the animate *iru*), even if it takes the nominative argument as its honorific target, is identified as a transitive verb when it expresses a possessive meaning.

(37) a. *Kimi$_i$ ni   mo   zibun$_i$ no   go-ryoosin   ga   irassyar-u   hazuda.*
          you   DAT   also self   GEN HON-parents NOM   have.ANM.HON PRS   should
          'You should also have your own parents.'

    b. *Ken   wa   kimi$_i$ ni   [PRO$_i$ rippana   go-ryoosin   ga   irassyat-te]*
          Ken   TOP you   DAT   excellent HON-parents NOM   have.ANM.HON GER

        hosikat-ta.
          want-PST
          'Ken wanted you to have excellent parents.'

    c. *[PRO_{arb} rippana go-ryoosin ga irassyaru] koto wa ii koto da.*
           excellent HON-parents nom have.ANM.HON that TOP good thing COP
   'It is a good thing PRO_{arb} to have excellent parents.'

Given that in (37), the dative possessor argument qualifies as a subject on all the subject tests, it is plausible to say that in the possessive construction, the dative argument counts as subject, which is a legitimate target for authentic subject honorification, and that in (33a) the nominative argument is exceptionally (or additionally) admitted as a target for honorification even though it is a non-subject argument (see Kishimoto 2004).

    As for the question of how subject honorification is triggered by a non-subject in (33a), Kishimoto (2000, 2012) suggests on the basis of the split vP hypothesis (see e.g. Chomsky 1995) that it becomes available by virtue of the shift of the nominative object to the specifier position of vP, which is the locus of animacy agreement.

(38) a. [_{TP} POSS-DAT [_{vP} ~~POSS-DAT~~ TH-NOM [_{VP} ~~TH-NOM~~ *iru*]]]

      b. [_{TP} POSS-DAT [_{vP} ~~POSS-DAT~~   [_{VP} TH-NOM *aru*]]]

      c. [_{TP} TH-NOM   [_{vP} ~~TH-NOM~~   [_{VP} ~~TH-NOM~~ *aru/iru*]]]

In Kishimoto's analysis, subject honorification targets an argument located within vP. The dative subject, base-generated in vP, is a canonical target of subject honorification, and if the nominative object is shifted to vP, it can be an additional target. The object shift is induced when *iru*, but not *aru*, is used as a possessive verb, as illustrated in (38a) and (38b). Consequently, the possessive *iru*, unlike the possessive *aru*, agrees obligatorily with the nominative argument, and at the same time allows its nominative object to be an additional target for subject honorification, as in (33a).

    The possessive *aru* and *iru* show an asymmetry in animacy agreement while the existential *aru* and *iru* do not. Kishimoto's analysis taking the nominative argument to trigger agreement when it is raised to vP also provides a ready account for the fact that both the existential *iru* and *aru* agree with their nominative argument. With the intransitive existential verbs, animacy agreement is forced, owing to the fact that the nominative argument (realized as a subject) is moved to vP via A-movement up to the clause-subject position of Spec-TP, as in (38c). The possessive *iru* takes a nominative object undergoing shift to vP, as in (38a), and hence must agree in animacy with its object. No agreement is forced on the possessive *aru*, since its nominative object is not moved to vP, as in (38b). Shibatani (1978) takes animacy agreement as a subject agreement, but the data discussed above suggest that it is triggered by an internal argument, which is not necessarily realized as a subject. Overall, both *aru* and *iru* have transitive possessive uses, contrary to the traditional claim.

4 Definiteness and transitivization

This section discusses the so-called "definiteness effect" observed for possessive predicates, and shows that the definiteness effect arises as a consequence of transitivization, which brings about the effect of turning an intransitive existential predicate into a transitive possessive predicate.

4.1 Possessive verbs and the definiteness restriction

It is commonly assumed that there is a close relationship between existential and possessive meanings (Jackendoff 1983; Pinker 1989). If we follow Kishimoto (2005) in assuming that *aru* and *iru* are originally intransitive, taking a theme argument as subject alongside a locative adjunct, it is plausible to postulate that possessive verbs are derived from existential verbs via transitivization (Heine 1997a, b), as illustrated in (39).[10]

(39) Lexical meaning: 'x is located at y'   'y possesses x'
      | | (REANALYSIS) | |
   Thematic role: Theme   Locative →   Possessor Theme
      | | | |
   Grammatical relation: SBJ   ADJUNCT   SBJ   OBJ

When an existential predicate acquires a possessive meaning, transitivization takes place whereby the original locative adjunct is reanalyzed as the subject of the possessive verb and the original theme argument as its direct object. Since unaccusative verbs do not normally take such a syntactic object, Kishimoto (2000) argues, the objects of the possessive verbs under reanalysis are subject to a special constraint, i.e. the definiteness restriction. This restriction, first observed by Milsark (1974, 1977), states that the definite class of DPs cannot appear in the English existential construction (e.g. *There were {these/all/most} boys in the room) while the indefinite class of DPs can (e.g. *There were {some/many/no} boys in the room*). It is shown below that the same definiteness restriction is indeed at work in the Japanese possessive constructions.

First, the nominative-marked theme argument in the possessive construction is constrained by the definiteness restriction regardless of whether the verb is *aru* or *iru*, as seen in (40).

---

10 The expletive *there* originates as a locative adjunct etymologically (see e.g. Huddleston and Pullum 2002), but in the existential *there* construction, *there* fills a subject position, and the theme appears in postverbal position, constrained by the definiteness restriction. The fact suggests that both Japanese possessive construction and English *there* construction are derived via the same grammatical process of transitivization.

(40) a. *Ken   ni   {takusan/huta-ri}   no   kyoodai   ga*
           Ken   DAT   {many/two-CLF}   GEN   brother   NOM

           {ar-u/i-ru}.
           {have.INAN PRS/have.ANM PRS}
           'Ken has {many/two} brothers.'

     b. **Ken   ni   {hotondo/subete/ryoohoo}   no   kyoodai   ga*
           Ken   DAT   {most/all/both}   GEN   brother   NOM

           {ar-u/i-ru}.
           {have.INAN PRS/have.ANM PRS}
           'Ken has {most/all/both} brothers.'

In contrast to (40a), where *takusan* 'many' and *huta-ri* 'two' are compatible with the nominative arguments of both *aru* and *iru* on their possessive meaning, (40b) is ungrammatical because the nominative arguments involve quantifiers like *subete* 'all', *hotondo* 'almost', and *ryoohoo* 'both'. The same peculiarity is observed for the post-verbal arguments of the English existential *there*-sentences (Milsark 1974, 1977).

(41) a. *There are {most/all/both} books on the table.

     b. There are {two/many/some} books on the table.

This parallelism is not too surprising because *aru* and *iru* are the Japanese counterparts of the existential verb *be* in English.

The definiteness restriction does not apply to the dative arguments of the possessive verbs, however, as seen in (42).

(42) a. *{Hotondo/Subete}   no   hito   ni   kyoodai   ga*
           {most/all}   GEN   person   DAT   brother   NOM

           {ar-u/i-ru}.
           {have.INAN PRS/have.ANM PRS}
           '{Most/All} people have brothers.'

     b. *{Takusan/Huta-ri}   no   hito   ni   kyoodai   ga*
           {many/two-CLF}   GEN   person   DAT   brother   NOM

           {ar-u/i-ru}.
           {have.INAN PRS/have.ANM PRS}
           '{Many/Two} people have brothers.'

Similarly, the definiteness effect is not observed for the arguments of the existential *aru* and *iru*. When *aru* and *iru* are used as existential verbs, neither the nominative argument nor the *ni*-marked locative adjunct is constrained by the definiteness restriction, as seen in (43).

(43) a. *Tana ni {hotondo/subete/takusan/ikuraka} no hon ga ar-u.*
　　　 shelf LOC {most /all/many/some} GEN book NOM be.INAN PRS
　　　 '{Most/All/Many/Some} books are on the shelf.'

　　 b. *Kooen ni {hotondo/subete/takusan/san-nin} no hito ga i-ru.*
　　　 park LOC {most/all/many/three-CLF} GEN person NOM be.ANM PRS
　　　 '{Most/All/Many/Three} people are in the park.'

　　 c. *{Hotondo/Subete/Takusan} no uti ni terebi ga ar-u.*
　　　 {most/all/many} GEN home LOC television NOM be.INAN PRS
　　　 'There is a television set at {most/all/many} homes.'

The data indicate that only the nominative arguments of the possessive *aru* and *iru* are constrained by the definiteness restriction.

　　 Just as noun phrases are divided into definite and indefinite classes, so too are *wh*-words partitioned into the two classes. The *wh*-words *dare* 'who' and *dono* 'which' count as definite, and *ikutu* 'how many' and *donna* 'what' as indefinite for the purpose of the definite restriction. Thus, the following contrast in acceptability is obtained.

(44) a. **Ken ni wa {dare no/dono} bessoo ga ar-u no?*
　　　 Ken DAT TOP {who GEN/which} villa NOM be.INAN PRS Q
　　　 '{Whose/Which} villa does Ken have?'

　　 b. *Ken ni wa {ikutu/donna} bessoo ga ar-u no?*
　　　 Ken DAT TOP {how.many/what} villa NOM be.INAN PRS Q
　　　 '{How many/What} villas does Ken have?'

The definiteness restriction constrains the possibility of relativization as well. In possessive sentences, the dative argument, but not the nominative argument, can be relativized.

(45) a. *?*[Ken ni t_i i-ru] kodomo$_i$*
　　　 Ken DAT 　 have.ANM PRS child
　　　 'the child who Ken has'

　　 b. *?*[Mari ni t_i ar-u] zaisan$_i$*
　　　 Mari DAT 　 have.INAN PRS property
　　　 'the property which Mari has'

　　 c. *[t_i kodomo ga i-ru] Ken$_i$*
　　　 　 child NOM have.ANM PRS Ken
　　　 'Ken, who has a child'

No such constraint obtains in existential sentences, since both the *ni*-marked locative adjunct and the nominative theme argument can be turned into the heads of relative clauses.

(46) a. *[syokki-dana   ni   t_i   ar-u]   koppu_i*
          cupboard   LOC   be.INAN PRS   cup
          'a cup that is on the cupboard'

      b. *[t_i   koppu   ga   ar-u]   syokki-dana_i*
              cup   NOM   have.INAN PRS   cupboard
          'the cupboard on which cups are placed'

Note that the nominative argument constrained by the definiteness restriction can be placed in cleft-focus position, as in (47a), and permits amount relativization, as in (47b).

(47) a. *[Ken-ni   t_i   ar-u]   no-wa   okane_i   (dake) da.*
         Ken DAT   have.INAN PRS that-TOP money   only   COP
         'What Ken has is money (only).'

     b. *Kare_i wa [zibun_i ni   ar-u]   subete no   zaisan o   nagedasi-ta.*
         he   TOP self   DAT be.INAN PRS all   GEN property ACC throw-PST
         'He threw away all the property that he had.'

In amount relatives, the head noun needs to occur with a modifier (or a determiner) like *subete* 'all' and *arayuru* 'every' (Carlson 1977).

   Let us now turn to the example in (48), on the basis of which Shibatani (1978) argues that the animate *iru* cannot have a transitive possessive use.

(48) *Abe-sensei   ni   sirami   ga   {i-ta/*o-ide-ni-nat-ta}.*
      Abe-teacher   LOC   louse   NOM   {be.ANM PST/be.ANM.HON PST}
      'There were lice on Professor Abe.'

Shibatani claims that since subject honorification cannot target the *ni*-marked argument in (48), the verb *iru* can only be used as an intransitive predicate. Nevertheless, the reason subject honorification fails in (48) is simply that this sentence is a bona fide existential clause, as confirmed by the fact that the nominative argument can undergo relativization.

(49) a. *[Abe-sensei   ni   t_i   i-ta]   sirami_i*
          Abe-teacher   LOC   be.ANM PST   louce
          'the lice, which were on Professor Abe'

b. ?*[Abe-sensei *ni*   *t_i*   *i-ta]*   *kodomo_i*
   Abe-teacher   DAT   be.ANM PST   child
   'a child, who Professor Abe had'

As shown in (49a), a sentence like (48) tolerates the relativization of the nominative argument, and thus crucially differs from an authentic possessive clause, which does not allow the nominative argument to undergo relativization, as in (49b).

In (48), the *ni*-marked adjunct happens to refer to a human, but it is identified as a location in thematic terms. Thus, the meaning does not change significantly even if a noun like *karada* 'body' is added to the *ni*-marked locative adjunct, as shown in (50a).

(50) a. *Ano   otoko   (no   karada)   ni   sirami   ga   i-ta.*
         that   man   (GEN   body)   LOC   louce   NOM   be.ANM PST
         'There were lice on that man('s body).'

      b. *Ano   otoko   (*no   karada)   ni   mago   ga   i-ta.*
         that   man   (GEN   body)   LOC   grandchild   NOM   be.ANM PST
         '{That man/*That man's body} had a grandchild.'

On the other hand, the possessor must be human unless a part-whole relation is expressed. Thus, in a possessive sentence like (50b), it is not possible to add a noun like *karada* to the dative argument. If (48) is an existential clause, there is a sense in which subject honorification should not be triggered by the *ni*-marked phrase, which is a locative adjunct.

Regular transitive stative predicates like *wakaru* 'understand' taking the <DAT/NOM, NOM> case frames do not have the distributions observed above for the possessive verbs. In particular, the second nominative argument of such a stative predicate is not constrained by the definiteness restriction. Thus, both definite and indefinite classes of DPs can appear as the nominative argument, and this argument can undergo relativization, as seen in (51).

(51) a. *Ken   ni   wa   {subete/hotondo/takusan}   no   kotoba*
         Ken   DAT   TOP   {all/most/many}   GEN   language

         ga   wakar-u.
         NOM   understand-PRS
         'Ken understands {all/most/many} languages.'

      b. *[Ken   ni   t_i   wakar-u]   kotoba_i*
         Ken   DAT   understand-PRS   language
         'the language which Ken understands'

In addition, subject honorification can be triggered by the dative subject, but not the nominative object, as illustrated in (52).

(52) a. *Abe-sensei   ni   wa   Mari   ga   o-mie-ni-nar-u.*
Abe-teacher DAT TOP Mari NOM can.see-HON PRS
'Professor Abe can see Mari.'

b. **Ken   ni   wa   Abe-sensei   ga   o-mie-ni-nar-u.*
Ken DAT TOP Abe-teacher NOM can.see-HON PRS
'Ken can see Professor Abe.'

In ordinary dative-subject constructions, the nominative object is never targeted for subject honorification. This fact follows, given that the nominative object does not undergo object shift to vP, due to the lack of animacy agreement, as Kishimoto (2012) discusses (see Section 3).

Observe at this point that while the existential/possessive *aru* is most typically used as a stative predicate, it can also be used as a non-stative predicate while carrying the meaning of 'an event x takes place', as in (53).

(53) *Kono   kaigi   ni   wa   takusan   no   sankasya   ga   ar-u.*
this meeting LOC TOP many GEN participant NOM be.INAN PRS
'There will be a lot of participants at this meeting.'

In (53), *aru*, which appears in the present form, refers to a future event, and the sentence expresses an eventive meaning. The verb *aru* can be used in an eventive sense when the nominative argument is an event noun (e.g. *ziko* 'accident'), and this argument could also be a human noun (e.g. *sankasya* 'participant') referring to an individual taking part in an event (Kageyama 2004). Note that no animacy agreement obtains here, as can be seen by the fact that the nominative argument can be animate.

Kishimoto (2000) suggests that in a sentence like (54), while the verb *aru* itself is a stative verb, an eventive meaning emerges when it is combined with a neo-Davidsonian "event argument", which is inserted into the subject position.

(54)   [TP EVENT-ARG [vP [VP TH-NOM *aru*]]]

On the assumption that the eventive *aru* is transitivized by virtue of an event argument, despite the fact that it superficially has only one argument (plus a locative adjunct), it naturally follows that the nominative argument behaves as an object, which does not trigger animacy agreement, as in (53). Furthermore, given the configuration in (54), it is predicted that the nominative argument will be constrained by the definiteness restriction. The nominative argument is indeed subject to this

constraint, as confirmed by the contrast in acceptability in (55) arising from the type of DP serving as the nominative theme argument.

(55) a. **Kono kaigi   ni   {hotondo/subete} no   sankasya   ga   ar-u.*
         this   meeting LOC {most/all}   GEN participant NOM be.INAN PRS
         'There will be {most/all} participants at this meeting.'

   b. *Kono kaigi   ni   {takusan/nan-nin-ka} no   sankasya   ga   ar-u.*
         this   meeting LOC {many/some-CLF Q}   GEN participant NOM be.INAN PRS
         'There will be {many/some} participants at this meeting.'

The fact that the nominative argument cannot be relativized, shown in (56), also illustrates that it is constrained by the definiteness restriction.

(56) a. **[kaigi   ni t_i ar-u]*   *(takusan   no)   sankasya$_i$*
         meeting   LOC   be.INAN PRS (many   GEN) participant
         '(many) participants, who will be at the meeting'

   b. *[t_i takusan   no   sankasya   ga   ar-u]*   *kaigi$_i$*
         many   GEN participant NOM be.INAN PRS meeting
         'the meeting which will have many participants'

What is more, the nominative argument of the eventive *aru* fails to trigger subject honorification, but subject honorification targeting the nominative argument is allowed when the verb has an existential use, as shown in (57).

(57) a. **Kono kaigi   ni   sankasya   ga   o-ari-ni-nar-u.*
         this   meeting LOC participant NOM be.INAN.HON PRS
         'There will be participants in this meeting.'

   b. *Asoko ni   kaigi   no   sankasya   ga   irassyar-u.*
         there LOC meeting GEN participant NOM be.ANM.HON PRS
         'The participants of the meeting are there.'

Given that these properties are shared by the transitive possessive *aru*, it is reasonable to analyze the eventive *aru* in (53) as a predicate taking two arguments, one of which is a hidden "event" argument.

4.2 Transitivization of eventive unaccusative verbs

In addition to the existential verbs *aru* and *iru*, some non-stative unaccusative verbs with eventive meanings are subject to transitivization when they express the meaning of 'coming into possession'. Representative examples include the intransitive

verbs *umareru* 'be born', *dekiru* 'be made, come out', and *tanzyoo-suru* 'be born', which can accompany *ni*-marked locatives (Kishimoto 2005; Masuoka 2000).

(58) a. *Tanaka-san no uti ni akatyan ga umare-ta.*[11]
      Tanaka-Mr. GEN home LOC baby NOM be.born-PST
      'A baby was born at Mr. Tanaka's home.'

    b. *Kobe ni siten ga deki-ta.*
      Kobe LOC branch NOM be.made-PST
      'A branch office was opened in Kobe.'

When the verbs are used intransitively, the nominative arguments are subjects. Accordingly, they can serve as the antecedent of the reflexive *zibun* and can also be targeted for subject honorification.

(59) a. *Eri$_i$ ga zibun$_i$ no uti de umare-ta.*
      Eri NOM self GEN home in be.born-PST
      'Eri was born in her own home.'

    b. *Itoo-sensei ga sentyuu ni o-umare-ni-nat-ta.*
      Itō-teacher NOM war.middle in be.born-HON PST
      'Professor Itō was born during the war.'

When the verbs describe the event of establishing the relation of inalienable possession (or kinship), as exemplified in (60), the human argument in the dative is identified as a possessor rather than a locative.

(60) *Ken ni kodomo ga {umare-ta/deki-ta}.*
    Ken DAT child NOM {be.born-PST/be.made-PST}
    'A child was born to Ken.'

In such cases, the dative argument, i.e. the possessor, serves as subject, thus functioning as the antecedent of the reflexive *zibun* 'self' or as the target of subject honorification, as illustrated in (61).

(61) a. *Ken$_i$ ni (mo) zibun$_i$ no mago ga {umare-ta/deki-ta}.*
      Ken DAT (also) self GEN grandchild NOM {be.born-PST/be.made-PST}
      'His own grandchild was (also) born to Ken.'

---

11 The verb *umareru* 'give birth' can take a *de* marked locative as well. When the verb takes a *ni* marked locative, a possessive meaning is implied.

b. *Abe-sensei   ni   kodomo ga   {o-umare-ni-nat-ta/o-deki-ni-nat-ta}*.
   Abe-teacher DAT child   NOM {be.born-HON PST/be.made-HON PST}
   'A child was born to Professor Abe.'

The data suggest that the eventive verbs expressing the meaning of 'x comes to be located at y' should be reanalyzed as transitive predicates with the meaning of 'y comes into possession of x' by the same process of transitivization that changes the intransitive existential *aru/iru* to the transitive possessive *aru/iru*.

(62) Lexical Meaning: 'x comes to be located at y'   'y comes to possess x'
          | | (REANALYSIS) | |
     Thematic role: Theme   Locative   →   Possessor   Theme
          | | | |
     Grammatical
     relation: SBJ   ADJUNCT   SBJ   OBJ

It is important to see that the presence or absence of the component of meaning indicating an instantiation of a change, represented as 'come to' in (62), does not affect the applicability of transitivization. Both stative and non-stative verbs can be transitivized if a possessive meaning is encoded and the locative is reinterpreted as a possessor. The fact that some eventive verbs with the meaning of 'coming into possession' are derived via transitivization is not too surprising, given that English non-stative verbs like *begin* can appear in *there*-constructions, as in *There began a riot*, when they express the meaning of 'coming into existence' (McCawley 1988).

   If the intransitive verb *umareru* taking the frame <LOC, NOM> is turned into the transitive *umareru* with the frame <DAT, NOM> via the transitivization process in (62), it will be expected that the objects of the transitive *umareru* will be constrained by the definiteness restriction. This expectation is fulfilled, because the definite class of DPs is prevented from occurring as the nominative argument, as shown in (63).

(63) a. ?**Ken   ni   {subete/ryoohoo}   no   mago   ga   umare-ta*.
         Ken   DAT {all/both}   GEN grandchild NOM be.born-PST
         '{All/Both} grandchildren were born to Ken.'

   b. *Ken   ni   {huta-ri/takusan}   no   mago   ga   umare-ta*.
      Ken   DAT {two CLF/many}   GEN grandchild NOM be.born-PST
      '{Two/Many} grandchildren were born to Ken.'

Besides, the transitive *umareru* does not allow its nominative argument to be relativized, but this argument can be dislocated by pseudo-clefting, as in (64a, b). No problem arises with the relativization of the dative argument, as in (64c).

(64) a. ?*[Ken ni t_i umare-ta]   kodomo$_i$
   Ken   DAT   be.born-PST   child
   'the children, who were born to Ken'

  b. [Kyonen Ken   ni t_i umare-ta   no]   wa   onnanoko$_i$ da.
    last.year Ken   DAT   be.born-PST that TOP girl   COP
    'It was a girl that was born to Ken last year.'

  c. [t_i mago   ga   umare-ta]   Ken$_i$
    grandchild NOM be.born-PST Ken
    'Ken, to whom the grandchildren were born'

This type of restriction is not observed when *umareru* is used intransitively. Thus, (65a) is acceptable regardless of whether the nominative phrase falls into the definite or indefinite class, and it also allows relativization, as in (65b).

(65) a. *Kyonen {subete/ryoohoo/huta-ri/takusan} no   mago   ga   umare-ta.*
   last.year {all/both/two-CLF/many}   GEN grandchild NOM be.born-PST
   '{All/Both/Two/Many} grandchildren were born last year.'

  b. *[kyonen t_i umare-ta]   kodomo$_i$*
    last.year   be.born-PST child
    'a child, who was born last year'

The data show that with eventive possessive verbs like *umareru* and *dekiru*, the nominative arguments, realized as objects, display the same definiteness effect that is observed for the nominative arguments of the possessive verbs *aru* and *iru*.

  Incidentally, the verb *dekiru* 'can do' in (66) – the potential form of the verb *suru* 'do' – takes the case pattern of <DAT/NOM, NOM>, in the same way that the possessive verb *dekiru* 'be made' in (60) does.

(66) *Ken   ni   wa   hukuzatuna keisan   ga   deki-ru.*
  Ken   DAT TOP complex   calculation NOM can.do-PRS
  'Ken can make complex calculations.'

With the potential *dekiru* 'can do', just like the possessive *dekiru*, the initial dative argument can be the antecedent of the reflexive *zibun*, showing that the verb is transitive.

(67) *Ken$_i$ ni   (mo)   zibun$_i$ no   keisan   ga   deki-ru.*
  Ken   DAT (also) self   GEN calculation NOM can.do-PRS
  'Ken can (also) make his own calculation.'

The possessive verb *dekiru* 'be made' is not a transitive predicate *per se*, because it is derived from its intransitive counterpart by way of transitivization. By contrast, the potential *dekiru* is a bona fide transitive stative predicate. Thus, the potential *dekiru* does not impose the definiteness restriction on its nominative argument, as shown in (68).

(68) a. *Ken   ni   wa   {subete/hotondo/takusan/ikuraka}   no   keisan*
         Ken   DAT   TOP   {all/most/many/some}   GEN   calculation

         ga   deki-ru.
         NOM   can.do-PRS
         'Ken can make {all/most/many/some} calculations.'

     b. *[Ken   ni   t_i   deki-ru]   keisan_i*
         Ken   DAT   can.do-PRS   calculation
         'the calculation that Ken can make'

To recapitulate, the transitive stative predicates taking <DAT/NOM, NOM> case-marking patterns refer to present states when they appear in the present (or non-past) from. In contrast, verbs like *umareru* 'be born' and *tanzyoo-suru* 'be born' refer to future events when they appear in the present form. The latter group of verbs does not fall into the class of authentic stative verbs. Nevertheless, they behave as transitive predicates when used as verbs of coming into possession, because they are originally intransitive, but are transitivized to take two arguments. Remarkably, when transitivization applies, the object of a "non-stative" verb may be marked with nominative case.

5 List possessives and information structure

As previously discussed, the nominative arguments of the possessive verbs *aru* and *iru* are constrained by the definiteness restriction. Nevertheless, we still can find cases where this restriction apparently fails to apply, as exemplified in (69).

(69) a. *Ken   ni   wa   Mari   ga   i-ru.*
         Ken   DAT   TOP   Mari   NOM   have.ANM PRS
         'Ken has Mari.'

     b. *Ken   ni   wa   ano   kuruma   ga   ar-u.*
         Ken   DAT   TOP   that   car   NOM   have.INAN PRS
         'Ken has that car.'

The sentences in (69) are fully acceptable, despite the fact that the nominative arguments are definite. This gives us the impression that the nominative arguments

in (69) are exempt from the definiteness restriction, but this is in fact not the case. Following Kishimoto (2005), it is shown below that the definite expressions are admitted for the nominative arguments in (69) on the grounds that the sentences have so-called "list interpretations".

To make the point clearer, recall that the definiteness restriction applies to the post-verbal argument of the English existential *there*-construction; nevertheless, there are cases where it looks as if this definiteness restriction does not apply, as in (70).

(70)   There's Mary, John, and Susan.

Sentence (70) is acceptable because it has a "list" interpretation, where the sentence is used to enumerate individuals to identify a membership in response, for example, to a *wh*-question like *Who were at the party?* (Rando and Napoli 1978). (A list answer may consist of only a single member, and thus, *There's Mary.* could be a legitimate reply to the question.)

The Japanese examples in (69) share important properties with (70). While ordinary possessive sentences denote a possessive relation, including inalienable and alienable relations, the sentences in (69) are used to list possible candidates among a set of alternatives. The sentences express some pragmatically conceivable relation established between the two arguments (rather than a genuine possessive relation). For example, *Mari* in (69a) could refer to a person who just gives support to Ken, and *ano kuruma* 'that car' in (69b) could be a car which Ken uses but which does not belong to him. In such cases, the definiteness effect is apparently absent.

Given the apparent absence of the definiteness effect, one might suspect that the sentences in (69) do not qualify as transitive possessive sentences. Nevertheless, there is good reason to believe that these clauses are construed as variants of the possessive construction. For one thing, the impossibility of relativization in (71) suggests that the definiteness restriction is actually operative on the nominative arguments.

(71)   a. *[Ken   ni t_i i-ru]   *Mari$_i$*
           Ken   DAT   have.ANM PRS   Mari
           'Mari, who Ken has'

       b. *[Ken   ni t_i ar-u]   *ano   kuruma$_i$*
           Ken   DAT   have.INAN PRS   that   car
           'that car, which Ken has.'

Needless to say, relativization is applicable to the dative argument *Ken ni*, which is not constrained by the definiteness restriction, as shown in (72).

(72) a. *[t$_i$ Mari ga i-ru] Ken$_i$*
        Mari NOM have.ANM PRS Ken
      'Ken, who has Mari'

    b. *[t$_i$ ano kuruma ga ar-u] Ken$_i$*
        that car   NOM have.INAN PRS Ken
      'Ken, who has that car'

Furthermore, in a sentence like (69a), when the verb *iru* is used, subject honorification may be triggered by either the dative or the nominative phrase, as in (73) (see Section 3).

(73) a. *Abe-sensei ni wa Kyoko-san ga irassyar-u.*
        Abe-teacher DAT TOP Kyoko-Ms. NOM have.ANM.HON PRS
      'Professor Abe has Kyoko.'

    b. *Masao-kun ni wa Abe-sensei ga irassyar-u.*
        Masao-Mr. DAT TOP Abe-teacher NOM have.ANM.HON PRS
      'Masao has Professor Abe.'

Overall, the sentences in (69) display exactly the same syntactic behavior as authentic possessive clauses with regard to relativization and subject honorification, even though their nominative phrases accommodate definite expressions.

Ward and Birner (1995) argue that the post-verbal nominal in a list sentence like (70) instantiates a variable (representing hearer-new information) while the rest of the sentence serves as an open proposition (constituting a presupposition component), and that when this information structure obtains, definite expressions are allowed in the post-verbal position (provided they have hearer-new discourse information status).[12] This account can be applied to the Japanese examples in (69), which are identified as list possessives used to identify individuals to fit into the descriptions among a set of individuals to choose from. If a list possessive contains a variable that represents a hearer-new entity, it is plausible to state that the apparent exception to the definiteness restriction in (69) is attained via the formation of information structure where the nominative phrase is focused and the rest constitutes presupposed or background information.

Under this analysis, the sentences in (69), unlike ordinary possessive sentences, have information structure similar to that of pseudo-cleft sentences, where a focus is structurally separated from the presupposed part. There are a number of ways to confirm the differences in the information status between the two types of possessive clauses. One difference is found in the applicability of topicalization.

---

12 Many different proposals have been advanced in the literature in an attempt to account for when a definite expression can be introduced in the position constrained by the definiteness constraint.

(74) a. *Mari wa_i Ken ni   (mo)   t_i i-ru.
        Mari TOP Ken DAT (also)   have.ANM PRS
        'As for Mari, Ken (also) has one.'

    b. Kyoodai wa_i Ken ni   (mo)   t_i i-ru.
        brother TOP Ken DAT (also)   have.ANM PRS
        'As for a brother, Ken (also) has one.'

A topic codes old information shared by the hearer and the speaker. In the list pos-sessive sentence in (74a), the nominative argument refers to a hearer-new entity and thus resists topicalization, which encodes old information. In an ordinary possessive sentence like (74b), where no partitioning of focus and presupposition is imposed syntactically, the nominative argument can be topicalized because no conflict in information status ensues.

    Moreover, observe that while the dative phrase of a list possessive cannot be placed in cleft-focus position via pseudo-clefting, no problem arises with the place-ment of the dative phrase of an ordinary possessive sentence in cleft-focus position, as in (75).

(75) a. ?*[t_i Mari ga   i-ru   no] wa Ken_i (dake) da.
          Mari NOM have.ANM PRS that TOP Ken (only) COP
        'It is (only) Ken that has Mari.'

    b. [Kyoodai ga   t_i i-ru   no] wa Ken_i (dake) da.
        brother NOM   have.ANM PRS that TOP Ken (only) COP
        'It was (only) Ken that has a brother.'

The reason (75a) is unacceptable is that the dative phrase having old information status cannot be placed in focus position, which represents new information. On the other hand, (75b) is acceptable, because the dative phrase is not construed as being presupposed in itself. When the nominative phrase is placed in cleft-focus position, no difference in acceptability is observed, as seen in (76).

(76) a. [Ken ni   t_i i-ru   no] wa Mari_i (dake) da.
        Ken DAT   have.ANM PRS that TOP Mari (only) COP
        'It was (only) Mari that Ken has.'

    b. [Ken ni   t_i i-ru   no] wa otooto_i (dake) da.
        Ken DAT   have.ANM PRS that TOP brother (only) COP
        'It was (only) a younger brother that Ken has.'

In (76a), the nominative phrase represents new information, and in (76b), it is neutral with regard to the focus/presupposition partition. Both sentences are accept-

able, for no conflict in information status arises from placing the nominative phrase in cleft-focus position.

The two types of possessive clauses show yet another difference in regard to *wh*-questioning. With a list possessive, unlike an ordinary possessive, the dative phrase cannot be *wh*-questioned, because it counts as a presupposed element.

(77) a. ?*Dare   *ni*   *Mari*   *ga*   *i-ru*   *no?*
          who   DAT   Mari   NOM   have.ANM PRS   Q
          'Who has Mari?'

   b. *Dare*   *ni*   *kyoodai*   *ga*   *i-ru*   *no?*
         who   DAT   brother   NOM   have.ANM PRS   Q
         'Who has a brother?'

Note that *wh*-questioning asking for information on a presupposed constituent is generally not possible. This is confirmed by the pseudo-cleft sentence in (78).

(78) **[Nani*   *o*   *yon-da*   *no]*   *ga*   *Masao*   *na*   *no?*
         what   ACC   read-PST   that   NOM   Masao   COP   Q
         (lit.) 'It was Masao that read what?'

Like (78), (77a) is unacceptable because a presupposed component is *wh*-questioned.[13] No such *wh*-questioning constraint is imposed on an ordinary possessive sentence, since a focus/presupposition division is not formed syntactically.

In a list possessive sentence, the nominative phrase, which is construed as receiving focus, is not precluded from being turned into a *wh*-phrase, just as in the case of a nominative phrase of an ordinary possessive clause, as shown in (79).

(79) a. *Masao*   *ni*   *wa*   *(haigo*   *ni)*   *dare*   *ga*   *i-ru*   *no?*
          Masao   DAT   TOP   behind   LOC   who   NOM   have.ANM PRS   Q
          'Who does Masao have (behind him)?'

   b. *Masao*   *ni*   *wa*   *donna*   *kyoodai*   *ga*   *i-ru*   *no?*
         Masao   DAT   TOP   what   brother   NOM   have.ANM PRS   Q
         'What brother does Masao have?'

The *wh*-word *dare* 'who' counts as definite for the purpose of the definiteness restriction. Thus, (79a) can be uttered only as a list possessive sentence where a list of individuals to choose from is available in a context, showing that the nominative phrase referring to a discourse-new entity can be *wh*-questioned.

---

13 (77a), as well as (78), cannot be interpreted as an ordinary *wh* question asking for the value (or identity) of the *wh* phrase, but a kind of echo interpretation, which asks the hearer to provide the information that the speaker should be familiar with but has somehow missed, is still available.

There are some more peculiar properties observed for list possessive sentences, two of which I will discuss below. First, the negative form of a list possessive sentence in (80a) is not acceptable, in opposition to a regular possessive clause like (80b).

(80) a. **Ken   ni   wa   Mari   ga   i-na-i.*
        Ken   DAT TOP Mari   NOM   have.ANM NEG PRS
        'Ken does not have Mari.'

    b. *Ken   ni   wa   kyoodai   ga   i-na-i.*
        Ken   DAT TOP brother   NOM   have.ANM NEG PRS
        'Ken does not have a brother.'

    c. *Ano   toki   Ken   ni   wa   Mari   ga   i-nakat-ta.*
        that   time   Ken   DAT TOP Mari   NOM   have.ANM NEG PST
        'At that time, Ken did not have Mari.'

The negative form of a list possessive sentence is acceptable in some contexts, however; (80c) is judged to be acceptable when the speaker, holding the positive assumption that Mari is related to Ken in the present, asserts contrastively that the relationship does not hold in the past. Apparently, this comes from a semantic/ pragmatic constraint.

Secondly, list possessive clauses are not possible, or odd at best, if accompanied by modal expressions indicating a lack of certainty.

(81) a. *Ken ni   wa Mari ga   i-ru   {hazuda/sooda/??kamosirenai}.*
      Ken DAT TOP Mari NOM have.ANM PRS {should/seem/might}
      'Ken {should/seems to/might} have Mari.'

    b. *Ken ni   wa okusan ga   i-ru   {hazuda/sooda/kamosirenai}.*
      Ken DAT TOP wife   NOM have.ANM PRS {should/seem/might}
      'Ken {should/seems to/might} have a wife.'

In (81a), the difference in acceptability is contingent upon the speaker's belief about the proposition. The modal *sooda* 'seem, I hear' or *hazuda* 'should' indicates that the speaker is to some extent certain about the accuracy of information. In contrast, *kamosirenai* 'might' indicates that the speaker does not have confidence as to whether the described situation holds in the actual world. The deterioration of the list possessive sentence is observed if it occurs with *kamosirenai*. In contrast, no deterioration is observed for an ordinary possessive sentence in (81b) regardless of the choice of modal.

In a nutshell, some behavioral differences observed between ordinary possessive and list possessive sentences come from their information status. Ordinary possessive sentences have no syntactic partitioning of focus and presupposition, whereas

list possessive sentences have a special information structure in which the dative phrase constitutes a presupposed component and the nominative phrase a focused constituent.

6 Conclusion and future research perspectives

This chapter has discussed transitivity issues surrounding stative and existential/possessive verbs, which arise mainly from the fact that stative predicates mark their arguments differently from non-stative verbs. Perception/possessive/necessity verbs carrying the meaning of 'x is in the state of possessing/perceiving/needing y' are two-place stative predicates and are identified as transitive with the initial experiencer/possessor arguments serving as subjects and the second theme arguments as objects. The transitivity of dyadic stative predicates is diagnosed by a number of tests for subjecthood and objecthood.

Among these transitive stative predicates, possessive predicates carrying the meaning of 'x (comes to) possess y' are derived by reanalyzing the locative adjuncts of existential predicates into possessors. Existential predicates, which carry the meaning of 'x is (*or* comes to be) located at y', are intransitive, in that the theme argument x is realized as a subject, and the locative argument as a *ni*-marked adjunct. On the other hand, possessive predicates are transitive, where the theme argument serves as an object marked with nominative case, with the possessor being realized as a dative subject, owing to transitivization. Although nominative marking is generally not available for objects of non-stative predicates, non-stative possessive predicates can mark their objects with nominative case. Obviously, this state of affairs emerges from the fact that the locative adjunct of an existential predicate, whether it carries a stative or non-stative meaning, can be reanalyzed as a possessor through transitivization, insofar as the sentence is identified as expressing a possessive relation.

Some special properties are observed for transitivized possessive predicates. The nominative argument of a derived possessive verb shows peculiar behavior, constrained by the definiteness restriction. Nevertheless, there are also cases where this restriction apparently does not apply to the nominative argument. Apparent exceptions are observed when the possessive clauses have list interpretations.

There are many facets of stative predicates that have not been discussed in this chapter but merit extensive research. One such issue is the question of why stative predicates show the peculiar case-marking patterns. This has to with a strong cross-linguistic correlation between stativity and non-canonical case marking, dative subjects in particular, and perhaps, also with the fact that stative verbs tend to appear in impersonal constructions cross-linguistically. Another issue concerns the fact that stative predicates are largely divided into two sub-groups. One class consists of

experiencer-subject predicates, which include potential predicates, and the other class the existential and possessive verbs. The two groups of stative predicates possess distinct properties and hence sometimes show distinct behavior, even though they show similar surface case-marking patterns. In this respect, existential/possessive predicates seem to represent a marked case, and experiencer-predicates an unmarked case. Arguably, this holds true cross-linguistically. If so, the question arises as to why existential/possessive predicates should be distinguished from experiencer-subject predicates in this way.

Acknowledgments

References

Carlson, Greg. 1977. Amount relatives. *Language* 53. 520–542.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Haig, John. 1981. Are traversal objects objects? *Papers in Linguistics* 14. 69–101.

Harada, S.I. 1976. Honorifics. In Masayoshi Shibatani (ed.), *Syntax and semantics* 5: *Japanese generative grammar*, 499–561. New York: Academic Press.

Hashimoto, Shinkichi. 1969. *Joshi jodōshi no kenkyū* [Studies on particles and auxiliaries]. Tokyo: Iwanami Shoten.

Heine, Bernd. 1997a. *Cognitive foundations of grammar*. New York: Oxford University Press.

Heine, Bernd. 1997b. *Possession: Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Iori, Isao. 1995. Ga sitai to o sitai: Kakuhyōji no yure ni kansuru itikōsatsu [Ga sitai and o sitai – A note on case-alternation in Japanese]. *Nihongokyōiku* 86. 52–64.

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jacobsen, Wesley. this volume. Chapter 15: Lexical meaning and temporal aspect in Japanese. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kageyama, Taro. 2004. Sonzai-shoyū no keidōshi-kōbun to imi-hen'nyū [Existential-Possessive light verb constructions and semantic incorporation]. In Taro Kageyama and Hideki Kishimoto (eds.), *Nihongo no bunseki to gengo ruikei* [An analysis of Japanese and language typology], 3–23. Tokyo: Kurosio Publishers.

Kinsui, Satoshi. 2006. *Nihongo sonzai hyōgen no rekishi* [A history of Japanese existential expressions]. Tokyo: Hituzi Syobo.

Kishimoto, Hideki. 2000. Locational verbs, agreement, and object shift in Japanese. *The Linguistic Review* 17. 53–109.

Kishimoto, Hideki. 2004. Transitivity of ergative case-marking predicates in Japanese. *Studies in Language* 18. 105–136.

Kishimoto, Hideki. 2005. *Tōgokōzō to bunpō kankei* [Syntactic structures and grammatical relations]. Tokyo: Kurosio Publishers.

Kishimoto, Hideki. 2010. Subjects and constituent structure in Japanese. *Linguistics* 48. 629–670.

Kishimoto, Hideki. 2012. Subject honorification and the position of subjects in Japanese. *Journal of East Asian Linguistics* 21. 1–41.

Kishimoto, Hideki. forthcoming. Valency and case alternations in Japanese. In Taro Kageyama and Wesley Jacobsen (eds.), *Transitivity and valency alternation: Studies on Japanese and beyond*. Berlin: De Gruyter Mouton.

Kishimoto, Hideki and Satoshi Uehara. this volume. Chapter 2: Lexical categories. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kitahara, Yasuo. 1981. *Nihongo no bunpō* [Japanese grammar]. Tokyo: Chūō Kōron.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu and Yuki Johnson. 2005. On the non-canonical double nominative construction in Japanese. *Studies in Language* 29. 285–328.

Kuroda, S.-Y. 1978. Case-marking, canonical sentence patterns, and counter equi in Japanese (A preliminary survey). In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics*, 30–51. Tokyo: Kaitakusha.

McCawley, James. 1988. *The syntactic phenomena of English*. Chicago: University of Chicago Press.

Masuoka, Takashi. 1987. *Meidai no bunpō* [The grammar of propositions]. Tokyo: Kurosio Publishers.

Masuoka, Takashi. 2000. *Nihongo bunpō no shosō* [Aspects of Japanese grammar]. Tokyo: Kurosio Publishers.

Matsushita, Daizaburō. 1928. *Kaisen hyōjun nihon bunpō* [Revised standard Japanese Grammar]. Tokyo: Chūbunkan.

Mikami, Akira. 1953. *Gendai gohō josetsu* [Prolegomena to modern Japanese grammar]. Tokyo: Tōe Shoin.

Milsark, Gary. 1974. *Existential sentences in English*. Cambridge, MA: MIT dissertation.

Milsark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3. 1–29.

Onoe, Keisuke. 1997–1998. Bunpō o kangaeru: Shugo 1–4 [Thinking about grammar: Subjects 1–4]. *Nihongogaku* 16(11). 91–97, 16(12). 88–94, 17(1). 87–94, 17(4). 96–103.

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Rando, Emily and Donna Jo Napoli. 1978. Definites in *there*-sentences. *Language* 54. 300–313.

Sasaguri, Junko. 2000. Meishiku no modariti toshite no *koto* [*Koto* as a modal element of noun phrases]. In Yukiko Alam Sasaki (ed.), *Gengogaku to nihongo kyōiku* [Linguistics and Japanese language education], 161–176. Tokyo: Kurosio Publishers.

Shibatani, Masayoshi. 1978. *Nihongo no bunseki* [An analysis of Japanese]. Tokyo: Taishukan Shoten.

Shibatani, Masayoshi. 1984. Kaku to bunpō-kankei [Case and grammatical relations]. *Gekkan Gengo* 13(3). 62–70.

Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Shibatani, Masayoshi. 2001. Non-canonical constructions in Japanese. In Alexandra Y. Aikhenvald, R.M.W. Dixon and Masayuki Onishi (eds.), *Non canonical marking of subjects and objects*, 307–354. Amsterdam: John Benjamins.

Takubo, Yukinori. 2010. *Nihongo no kōzō* [The structure of Japanese]. Tokyo: Kurosio Publishers.

Tokieda, Motoki. 1950. *Nihonbunpō kōgohen* [Colloquial Japanese grammar]. Tokyo: Iwanami Shoten.

Tsunoda, Tasaku. 2009. *Sekai no gengo to nihongo, kaiteiban* [Languages of the world and Japanese, revised edition]. Tokyo: Kurosio Publishers.

Ward, Gregory and Betty Birner. 1995. Definites and the English existential. *Language* 71. 722–742.

Watanabe, Minoru. 1971. *Kokugo kōbun ron* [On Japanese grammar]. Tokyo: Haniwa Shobō.

Yamada, Yoshio. 1936. *Nihon bunpōgaku gairon* [An introduction to Japanese grammar]. Tokyo: Hōbunkan.

Naoyuki Ono
17 Agent nominals

1 Introduction

This chapter is concerned with the semantic and morphosyntactic properties of agent nominals in Japanese. For the sake of the discussion to follow, I will assume that "agent nominals" refer to morphologically complex nouns that denote a person who performs a particular action or is actually involved in an event denoted by the base predicate, or who is engaged in a specific profession or a habitual activity. The morphological means to derive such nominals in Japanese are essentially the same as those in English, although productivity differs in individual types of derivation. The examples in (1) illustrate suffixation (cf. Eng. *consumer, actor*), those in (2) deverbal compounding (cf. *lookout, pickpocket*), and those in (3) V-to-N conversion (cf. *guard, cheat*).

(1) a. *ryokoo-sya*   [travel-person]   'traveller'
    b. *hanasi-te*   [speak-hand]   'speaker'

(2) a. *mono-tori*   [thing-steal]   'thief'
    b. *kane-moti*   [money-own]   'wealthy person'

(3) a. *mihari*   [watch]   'watchman, guard'
    b. *minarai*   [apprenticeship] 'trainee'

Agent-forming suffixes attach to a base which may be a VN (verbal noun), noun, infinitive verb (*ren'yō* verb), or bound morpheme, as exemplified in (1), where the suffixes *-sya* and *-te* attach to the VN *ryokoo* 'travel' and the native infinitive verb *hanasi* 'speak', respectively. Most of the agent-denoting suffixes including *-sya* 'person' come from the Sino-Japanese (S-J) stratum, while those coming from the native Japanese stratum, like *-te* lit. 'hand', are limited in number. As seen from the English glosses, the "suffixes" in these examples originate from lexical nouns that have particular meanings like 'person' or 'hand' (which was metonymically extended to represent a person). The agent nominals in (2) are derived via compounding of a noun and an infinitive verb. As observed by Sugioka (1986), deverbal compounds incorporating the base verb's complement noun in the structure [N-V$_N$]$_N$ can be interpreted as an agent who performs the action denoted by the corresponding verb phrase "to verb a noun". For example, *mono-tori* [thing-steal] in (2a), comprising the noun *mono* 'thing' and the infinitive (or nominalized) verb *tori* 'steal', denotes a person who steals something from another person. It appears that the formation of

agent nominals by deverbal compounding is not as productive as the agent suffixation illustrated in (1). Conversion of simplex verbs is another type of agent nominal formation. In (3), the nouns *mihari* and *minarai* originate from the compound verbs *mi-haru* 'look out, watch' and *mi-narau* 'follow someone's example, receive training', respectively. This type of agent nominal appears to be even less productive than the deverbal compounding type of (2). Of the three types of agent nominal formation, this chapter will primarily deal with suffixation of type (1); the reader is referred to Chapter 9 (Yumoto, this volume) for more details of V-to-N conversion and deverbal compounds based on conversion.

The goal of the present chapter is to illuminate the interesting properties of suffixed agent nominals that have to do with general issues of the interface between morphology, semantics, and syntax. For this purpose, nouns of the following kinds will be excluded from our discussion: (i) those nouns that denote persons in particular social positions, roles, or posts; (ii) those that denote particular kinship relations; and (iii) those that represent human biological properties such as gender or age. For example, the nouns in (4), though morphologically complex, are excluded because they do not refer to a particular action performed by the denoted human-being.

(4) a. *sya-tyoo*   [company-head]   'president of a company'

  b. *ai-te*   [opposite-hand]   'opponent, counterpart'

  c. *titi-oya*   [father-parent]   'father'

  d. *nihon-zin*   [Japan-person]   'Japanese citizen'

Simplex nouns like those in (5) that are not derived by suffixation, compounding, or conversion are also left out of consideration.

(5) a. *doroboo*   'thief'

  b. *isooroo*   'rent-free lodger, freeloader'

Borrowings from English are not included in our discussion either, because they are not decomposed into morphemes in Japanese.

(6) a. *pianisuto*   'pianist'

  b. *puroguramaa*   'programmer'

We also exclude fashionable suffixes that are attributed to the foreign stratum by folk etymology, like *-raa* (mimicry of the English *-er*) as in *mayo-raa* [mayonnaise-er] 'mayonnaise freak'.

This chapter aims to present an extensive survey of agent nominals and to investigate their semantic and morphosyntactic properties in light of theoretical findings uncovered by previous research. The organization of this chapter is as follows. Section 2 will survey the morphological properties of selected types of agent nominals. We will see that agent nominals in Japanese are produced via suffixation, compounding, and conversion. Section 3 is concerned with argument structure conditions on agent suffixes. Some researchers have argued for a constraint that the denoted entity of an agent nominal should correspond to an external argument of the base verb. We will see that this constraint does not hold in Japanese agent nominals. Section 4 will discuss a well-studied topic regarding agent nominals, i.e. event vs. non-event interpretations of agent nominals. It will be shown that the difference in the modality of event interpretation is sometimes morphologically marked in Japanese. Section 5 will focus on the parallelism between agent nominals and event nouns. We will see that agent nominals with an eventive interpretation license adjuncts specifying for time, location, and means within the noun phrase. Thus, they behave syntactically more like event nouns. Section 6 summarizes and concludes the discussion.

2 Morphology of agent nominals

This section will account for the three morphological mechanisms involved in the formation of agentive nominals (suffixation, compounding, and conversion), using some representative examples.

2.1 Suffixation

Suffixation is the most productive type of word formation that produce agent nominals. There is an abundant variety of agent-denoting suffixes in Japanese, mostly in the Sino-Japanese stratum. This subsection will provide a brief description of seven representative suffixes which are particularly productive.

2.1.1 *-sya* (者)

The suffix *-sya* meaning 'person' is perhaps the most productive Sino-Japanese nominal suffix that attaches to verbal nouns, nouns, and S-J morphemes. The derived nominal denotes a person who actually performs an action or is involved in an event expressed by the base (Sugimura 1986; Kageyama 1999), as the examples in (7) and (8) show. In (7) the suffix attaches to Sino-Japanese verbal nouns that are freestanding words, and in (8) to Sino-Japanese bound morphemes.

(7) a. *unten-sya*   [drive-person]   'driver'

   b. *sigan-sya*   [apply-person]   'applicant'

   c. *riyoo-sya*   [use-person]   'user'

(8) a. *gaku-sya*   [study-person]   'scholar'

   b. *tyo-sya*   [write-person]   'author'

   c. *si-sya*   [die-person]   'dead person'

All the nouns given in (7) and (8) are qualified as "agent nominals" in the sense defined in Section 1 because they denote a person who performs a particular action (7a, 7b, 7c, 8b) or is actually involved in an event denoted by the base predicate (8c), or who is engaged in a specific profession or a habitual activity (8a). In actuality, however, the suffix -*sya* can derive not only agent nominal as defined in this chapter but also other human nouns that do not particularly imply an agentive action. Sugimura (1986) thus points out examples like *zituryoku-sya* [power-person] 'a powerful figure', whose meaning may be represented as 'a person who owns X (*zituryoku* 'power, real ability, influence'), *seisyoku-sya* [holy.position-person] 'priest, clergyman', which may be paraphrased as 'a person who is in the position of X (*seisyoku* 'religious vocation'), and *dokusin-sya* [unmarried-person] 'unmarried person', which designates 'a person who is in the condition of X (*dokusin* 'being unmarried').[1] These nouns do not concern us here.

2.1.2 -*ka* (家)

The agent-denoting suffix -*ka*, originating from a Sino-Japanese lexical noun representing 'house' or 'family', is attached to Sino-Japanese VNs, nouns, or bound morphemes to denote a person who has professional expertise in a specific field (Sugimura 1986). Many of the agent nominals derived by this suffix represent a profession that requires special knowledge and skills (Kageyama 1999). Observe the difference in meaning between (9) and (10), where the nominals with -*sya* denote persons who are actually involved in an action, while those with -*ka* denote persons who are professional performers of the action (Sugimura 1986).

---

1 The suffix *sya* produces more non agent nominals such as *higai sya* (suffer.harm person) 'victim', *higi sya* (suspect person) 'suspect', *tentoo sya* (fall.down person) 'fallen down person', *husyoo sya* (injury person) 'injured person', *siboo sya* (death person) 'dead person'. We will return to discussion of those examples in Section 3.

(9) a. *tozan-sya*   [mountain.climbing-person]   'climber'

       b. *tozan-ka*   [mountain.climbing-house]   'professional climber'

(10) a. *satuei-sya* [photography-person] 'photographer'

       b. *syasin-ka* [photo-house]   'professional photographer'

When attached to noun bases, *-ka* forms profession names as in (11), or sometimes represents enthusiasts or persons who have strong disposition toward something as in (12) (Sugimura 1986).

(11) a. *syoosetu-ka* [novel-house]   '(professional) novelist'

       b. *manga-ka*   [comic-house]   '(professional) comic artist'

       c. *syuukyoo-ka* [religion-house]   'priest'

(12) a. *bisyoku-ka*   [gourmet-house]   'epicure'

       b. *dokuzetu-ka* [poisonous.tongue-house] 'blistering critic'

       c. *aiken-ka*   [love.dog-house]   'dog lover'

In light of the definition of agent nominals given in the previous section, we assume that the example in (11) are agent nominals but those in (12) are not. In addition, the suffix attaches to Sino-Japanese morphemes to create agent nominals such as *sak-ka* [create-house] 'creator, novelist' and *ga-ka* [picture-house] 'painter'. It should be noted, however, that the productivity of the *-ka* suffix in this usage is very limited.

2.1.3 *-te* (手)

As a native Japanese suffix originally meaning 'hand', the suffix *-te* productively attaches to native infinitive verbs to produce agent nominals that denote persons who conduct particular actions in actual situations or are engaged in a profession or a habitual activity.

(13) a. *hanasi-te* [speak-hand]   'speaker'

       b. *kai-te*   [buy-hand]   'buyer'

       c. *tukuri-te* [produce-hand] 'producer'

       d. *utai-te*   [sing-hand]   'singer'

Many of such *-te* suffixed nouns come in pairs of opposite words, as in *hanasi-te* 'speaker, addressor' vs. *kiki-te* 'hearer, addressee', *kai-te* 'buyer' vs. *uri-te* 'seller',

and *tukuri-te* 'producer, maker' vs. *tukai-te* 'user'. It might be possible to attribute the paired formation to the fact that the agentive meaning of *-te* is a metonymic extension of its original meaning of 'hand',[2] and that a person normally has two hands. There are cases, however, where this characterization does not hold. For example, *nitooryuu no tukai-te* 'a two-sword fencer' can be used without a paired counterpart. It appears that such non-paired examples are licensed by special emphasis placed on the excellence of the person's special talent or skill.

Sugioka (1992) and Ito and Sugioka (2002) note that when a *-te* nominal takes a complement, it implies that a person is/was actually involved in a particular action denoted by the verb (and not a profession or a habitual activity), as in the following examples.

(14)   a.   *syoohin   no   kai-te*
            goods   GEN   buy-hand
            'buyer of the goods'

       b.   *monogatari   no   katari-te*
            story   GEN   speak-hand
            'teller of the story'

       c.   *seihin   no   tukuri-te*
            product   GEN   make-hand
            'producer of the product'

We will return to the action interpretation of agent nominals and the selection of arguments in Section 4.

2.1.4 *-syu* (手)

The Sino-Japanese suffix *-syu*, originally meaning 'hand', is represented by the same Chinese character as the native suffix *-te* above. The nominals derived by this suffix denote a person who is engaged in a particular profession. It may attach to S-J morphemes as in (15a, b, c) or freestanding VNs as in (15d, e).

---

2 Note that *te* (hand) is used as a noun forming suffix that derives nominals denoting other types of things, e.g. *tot te* (hold hand) 'handle', *moti te* (hold hand) 'handle', *kime te* (decide hand) 'conclu sive factor', *iku te* (go hand) 'the direction that one is facing'. The variety of meanings that the suffix can represent emerges because *te* 'hand' is figuratively extended to denote many different entities such as a human being, a part of an object, a means, or direction, etc. The agent interpretation is one of them.

(15) a. *ka-syu*   [sing-hand]   'singer'

b. *ki-syu*   [ride-hand]   'jockey'

c. *zyo-syu*   [assist-hand]   'assistant'

d. *unten-syu*   [drive-hand]   'driver'

e. *syageki-syu*   [shoot-hand]   'shooter'

2.1.5 *-nin* (人)

The S-J suffix *-nin* generally derives agent nominals by attaching to verbal nouns, as shown in (16).

(16) a. *tuukoo-nin*   [pass-person]   'passerby'

b. *hosyoo-nin*   [guarantee-person]   'guarantor'

c. *soozoku-nin*   [succeed-person]   'successor'

When attached to morphemes other than verbal nouns, this suffix forms human nouns that do not count as agent nominals for the purpose of this chapter, as in (17).

(17) a. *aku-nin*   [evil-person]   'bad person, villain'

b. *zen-nin*   [good-person]   'good-natured person'

c. *syoo-nin*   [marketing-person]   'merchant'

d. *byoo-nin*   [sick-person]   'sick person'

2.1.6 *-nusi* (主)

-Nusi, literally meaning 'master', 'proprietor', or 'head of a group', is a native Japanese suffix that is combined with native infinitive verbs to produce agent nominals. It implies that the person is actually involved in the denoted action or event.

(18) a. *moti-nusi*   [own-master]   'owner'

b. *okuri-nusi*   [send-master]   'sender'

c. *kai-nusi*   [buy-master]   'buyer'

This suffix is also used to form non-agent nominals when it is combined with non-verbal bases, as in *zi-nusi* [land-owner] 'land owner' and *ya-nusi* [house-owner] 'house owner'.

2.1.7 *-kyaku* (客)

Kyaku is a S-J word meaning 'guest'. This word functions as an agent-denoting suffix when attached to VNs to derive nominals that represent a persons who is actually involved in an event designated by the base. A few examples are shown in (19).

(19) a. *ryokoo-kyaku*   [travel-guest]   'traveller'

      b. *hoomon-kyaku*   [visit-guest]   'visitor'

      c. *tozan-kyaku*   [mountain.climb-guest]   'mountain climber'

2.1.8 The role of Chinese characters

What makes the inventory of Japanese agent suffixes more complicated than that of the English counterparts is the fact that in some of the productive suffixes, one and the same Chinese character is applied to both native Japanese and Sino-Japanese suffixes. Table 1 below shows correspondences between native and S-J suffixes represented by the same Chinese characters. The correspondences result from the general practice of dual (or multiple) reading for a single Chinese character: Japanese reading (*kun-yomi* or 'semantic reading') and Chinese reading (*on-yomi* or 'sound-reading) (cf. Chapter 1 [Kageyama and Saito, this volume]).

Table 1: Correspondences of Chinese characters and agent-deriving suffixes

	-者	-手	-人	-主	-家	-客
Native Japanese	mono	te	hito	nusi	–	–
Sino-Japanese	sya	syu	nin	syu	ka	kyaku

In light of Table 1, one might be tempted to assume that the paired suffixes that share one and the same Chinese character (e.g. *-mono* and *-sya* represented by -者, *-te* and *-syu* represented by -手) are two variants or allomorphs of a single morpheme. This view might gain support from the alleged complementary distribution of suffixes in nominalization that follows from the general combinatory principle that native suffixes are congruent with native bases and Sino-Japanese suffixes with Sino-Japanese bases. The fact, however, is not so simple as to be captured solely by restrictions on lexical strata. Consider, for example, the deverbal compound VN *moositate* 'allegation'. Since this word is native Japanese, the stratum restriction alone would predict that it should be combined with *-mono* or *-te*, but the predicted forms *moositate-mono and *moositate-te are judged ungrammatical; instead, *moositate-sya*, a hybrid combination of native and Sino-Japanese elements, is actually

attested. This attested form is obviously sanctioned by the categorial restriction of *-sya*, which selects for VNs.

2.2 Compounds

Compound formation is another way of producing agent nominals. Kageyama (1982), Sugioka (1986), and Ito and Sugioka (2002) point out that some verbal compounds comprising an incorporated noun and a verbal base express, among other things, the sense of agent of action.[3] The [N V$_N$] compounds in (20) are some of the agent compounds where the incorporated noun is the complement to the nominalized verb.

(20) a. *uso-tuki*   [lie-tell]   'liar'

        b. *mono-kaki* [thing-write]   'writer'

        c. *sake-nomi* [*sake*-drink]   '(heavy) drinker'

        d. *kyaku-hiki* [customer-pull] 'tout'

        e. *hito-gorosi* [human-kill]   'murderer'

Sugioka (1992) and Ito and Sugioka (2002) note that, in general, agent compounds differ from agent nominals derived by the suffixes *-te*, *-nusi*, *-kyaku* in that the former do not imply an action interpretation that the latter suffixed nominals often have. Compound nominals such as in (2) and (20) do not refer to actual actions but to habitual activities, characteristic behavior, or distinct attributes. This point is crucial when we consider the semantics of agent nominals in later sections.

2.3 Conversion

There are a limited number of agent nominals derived by conversion, as exemplified in (21) as well as in (3).

(21) a. *suri*   [steal]   'pickpocket'

        b. *yopparai* [get drunk]   'drunken person'

        c. *tukisoi*   [attend, accompany] 'attendant'

        d. *siriai*   [get to know]   'acquaintance'

        e. *tobi.iri*   [participate]   'unexpected participant'

---

3 The other meanings that deverbal compounds express include Instrument, Result/Product, Place/Time, Act/Event, Nominal predicates (See Sugioka 1986: 79 80).

Morphologically, the noun in (21a) comes from a simplex verb while the other nouns all originate from V-V compound verbs. Semantically, some of them denote persons who are involved in an actual, temporary action, as in *yopparai* 'drunken person' and *tobi.iri* 'unexpected participant', or in a habitual activity or a profession, as in *suri* 'pickpocket', *minarai* 'trainee', and *tukisoi* 'attendant'. It appears that conversion of this type has little or no productivity.

This section has outlined the basic morphological properties of the word formation processes involving agent nominals in Japanese. In what follows, we will turn attention to their semantic and syntactic peculiarities, with particular focus on theoretical issues that have been debated in the literature on English agent nominals. Specifically, the major issues that have emerged from the previous studies on English agent nominals derived by the suffix *-er* (cf. Booij 1986; Rappaport Hovav and Levin 1992; Bauer 1993; Ryder 1999; Lieber 2004; and many other works) will be classified into three: (i) argument structure conditions on agent suffixes, (ii) eventive versus non-eventive interpretations of nominals, and (iii) parallels between agent nominals and action (event) nominals. We will take up these issues in turn in Sections 3, 4, and 5 and bring to light the characteristic properties of Japanese agent nominals as against the English counterparts.

3 Argument structure conditions on agent suffixes

It is well known that the English suffix *-er* exhibits a variety of meanings when attached to verb bases, such as human agents (e.g. *singer*, *instructor*), instruments and devices (e.g. *scanner*, *printer*), themes (e.g. *broiler*, *fryer*), experiencers (e.g. *hearer*, *admirer*), and stimuli (e.g. *appetizer*, *thriller*). The wide range of interpretations associated with *-er* nominals has been a central topic of lexical semantics and syntax in English. Among the various proposals, what appears to be the more promising is to assume that *-er* nominals are derived only from verbs that have external arguments and that these nominals always refer to the external argument (cf. Keyser and Roeper 1984; Levin and Rappaport 1988; Rappaport Hovav and Levin 1992; Alexiadou and Schäfer 2010; among others). In particular, Alexiadou and Schäfer (2010) posit the generalization in (22) as the "external argument generalization".

(22)   The External Argument Generalization
       Agent nominals may correspond to the external argument of the base verb
       irrespective of the semantic role that the verb assigns to it.[4]

---

4 Shimamura (2006) claims that only event *er* nominals obey the external argument generalization, but Alexiadou and Schäfer (2010) assume that the external argument generalization may be valid for event (episodic) and non event (dispositional) *er* nominals.

The external argument generalization predicts that the verbs that lack external argument cannot form *-er* nominals (Levin and Rappaport 1988). This prediction is borne out in (23), where unergative verbs (i.e. intransitive verbs that have external argument) can provide bases for well-formed *-er* nominals whereas unaccusative verbs (i.e. intransitive verbs that do not have external argument) cannot.

(23) a. unergative bases: jumper, runner, walker, climber, wiggler, rider, …

      b. unaccusative bases: *arriver , *disappearer, *appearer, *dier, *laster, *ender, *exister, *happener, *occurrer, …

While the behavior of English *-er* thus appears to be properly regulated in terms of the notion "external argument" in argument structure, Japanese agent suffixes cannot be analyzed in the same way. Thus, the Japanese *-sya* 'person', the most productive agent suffix, is not constrained by argument types, as it is capable of attaching not only to transitive VNs (24) and unergative VNs (25) but also to unaccusative VNs (26).

(24) Transitive bases
    sinsei-sya [apply-person] 'applicant', *hakai-sya* [destroy-person] 'destroyer', *zyuken-sya* [take.exam-person] 'person who takes an exam, examinee', *kyooiku-sya* [educate-person] 'educator', *syookai-sya* [introduce-person] 'introducer', *rikai-sya* [understand-person] 'sympathizer, supporter'

(25) Unergative bases
    zisatu-sya [commit.suicide-person] 'a suicide', *tooboo-sya* [escape-person] 'runaway, escapee', *roodoo-sya* [work-person] 'laborer', *tuukin-sya* [commute-person] 'commuter', *yuusyoo-sya* [win-person] 'winner'

(26) Unaccusative bases
    siboo-sya [pass.away-person] 'a dead person', *tentoo-sya* [topple-person] 'a person who topples down', *zyoohatu-sya* [disappear-person] 'a missing person', *sonzai-sya* [exist-person] 'a being, creature', *tootyaku-sya* [arrive-person] 'a person who arrives', *rakka-sya* [drop-person] 'a person who drops'

Evidence showing the unaccusative status of VNs like those appearing as the bases of agent nominals in (26) has been presented by various researchers in their discussion on light verb constructions (Miyagawa 1989; Tsujimura 1990; Kageyama 1993; Kobayashi 2004; see Chapter 12 [Miyamoto and Kishimoto, this volume] for more discussion). There is a general consensus among researchers that verbal nouns are divided into transitive, unergative, and unaccusative classes, as shown above in (24), (25), and (26). A critical diagnosis is to test whether a given VN can bear the

accusative marker or not. The transitive VN *soozi* 'clean' and the unergative intransitive VN *sanpo* 'take a walk' behave similarly in that they can be used either as accusative objects of the verb *suru* as in (27a) and (28a) or as part of complex verbs without the accusative case as in (27b) and (28b).

(27) a. *Taroo   wa   heya   no   soozi   o   si-teiru.*
       Taro   TOP   room   GEN   cleaning   ACC   do-ASP
       'Taro is cleaning the room.'

    b. *Taroo   wa   heya   o   soozi-si-teiru.*
       Taro   TOP   room   ACC   cleaning-do-ASP.
       'Taro is cleaning the room.'

(28) a. *Taroo   wa   kooen   de   sanpo   o   si-teiru.*
       Taro   TOP   park   IN   walk   ACC   do-ASP
       'Taro is taking a walk in the park.'

    b. *Taroo   wa   kooen   de   sanpo-si-teiru.*
       Taro   TOP   park   IN   walk-do-ASP
       'Taro is taking a walk in the park.'

Unaccusative VNs, on the other hand, do not allow the accusative marking on them, and hence can occur only in the incorporated structure. Compare the unacceptability of (29b) with the acceptability of (29a).

(29) a. *Titi   ga   kinoo   siboo-si-ta.*
       father   NOM   yesterday   pass.away-do-PST
       'My father passed away yesterday.'

    b. **Titi   ga   kinoo   siboo   o   si-ta.*
       father   NOM   yesterday   pass.away   ACC   do-PST

Since the verb *suru* in this case is a light verb (Grimshaw and Mester 1988), lacking an argument structure of its own, it must have an argument structure transferred to it from the VN. Thus, *suru* has the argument structure of the VN whose argument structure is transferred to it. In the case of unaccusative VN's such as those in (35), the argument structure lacks an external argument. Thus, according to Burzio's generalization (i.e. only those predicates that take an external argument can assign accusative case), it also lacks the ability to assign accusative case to its complement. This explains the ungrammaticality of accusative marking on unaccusative VNs as in (29b).

Having observed that a structural condition like the external argument generalization motivated for English fails to apply to the Japanese suffix *-sya*, we are now in

a position to discuss the nature of *-sya* and other Japanese suffixes denoting agents. As briefly mentioned in Section 2, most agent-denoting suffixes in Japanese originate from lexical nouns meaning 'person', 'hand', or 'house', which are grammaticalized as suffixes representing human beings by way of metonymy or metaphor. Because of their semantic uniqueness, Japanese agentive suffixes do not have the kind of ambiguity that the English suffix *-er* exhibits. Thus the majority of productive agent nominals in Japanese uniquely denote agents of action, as shown in (30).

(30)　a.　*hokoo-sya*　　　[walk-person]　　'pedestrian'
　　　　　unten-sya　　　　[drive-person]　　'driver'

　　　b.　*tuukoo-nin*　　　[pass-person]　　'passer-by'
　　　　　sasidasi-nin　　[address-person]　'addressor'

　　　c.　*utai-te*　　　　　[sing-hand]　　　'singer'
　　　　　yomi-te　　　　　[read-hand]　　　'reader'

　　　d.　*moti-nusi*　　　　[own-master]　　'owner'
　　　　　yatoi-nusi　　　[employ-master]　'employer'

　　　e.　*hoomon-kyaku*　　[visit-guest]　　'visitor'
　　　　　syukuhaku-kyaku　[stay-guest]　　'hotel guest'

Since the meanings of the Japanese agent suffixes, whether native or Sino-Japanese, are lexically fixed, they cannot freely denote other things than human beings. If one wishes to express a non-agent entity by using the same verbal base, it is necessary to employ an appropriate object-denoting suffix instead of the agent suffix. This results in pairs like those in (31) and (32), where agents are represented by agentive suffixes and objects (themes) by object-denoting suffixes.

(31)　a.　*otosi-nusi*　　[lose-master]　　'person who lost his/her property,
　　　　　　　　　　　　　　　　　　　　　owner of a lost property'

　　　b.　*otosi-mono*　　[lose-thing]　　'lost property'

(32)　a.　*haiki-sya*　　[dispose.of-person]　'person whose disposes of something'

　　　b.　*haiki-butu*　[dispose.of-thing]　'thing disposed of, waste matter'

　　　Sugimura (1986) and Kageyama (1999) brought to light the curious fact that in limited cases, nominals derived by *-sya* or some other agent suffixes denote a "patient", i.e. a person who is affected or acted upon by the action denoted by the base verb, instead of an agent or actor. Observe the examples in (33), where *-sya*, which expresses an agent or actor in normal cases as we saw in Section 2, is apparently associated with a "reverse" interpretation.

(33) a. *taiho-sya*   [arrest-person]   'arrestee'

  b. *koyoo-sya*   [employ-person]   'employee'

  c. *saiyoo-sya*   [hire-person]   'hired person, new hire'

  d. *kaiko-sya*   [dismiss-person]   'dismissed person'

  e. *siyoo-nin*   [use-person]   'employee, servant'

The base of the nominal *taiho-sya* [arrest-person] 'arrestee' in (33a) is the transitive VN *taiho* 'arrest', which takes an agent and a theme (patient). Since *-sya* is generally associated with an agent argument of its base VN as we saw earlier, it would naturally be expected that *taiho-sya* would mean 'a person who arrests somebody'. Contrary to expectation, however, it actually means 'a person who is arrested', namely the patient of the arresting event. Does this suggest that something like an invisible passive morpheme is hidden in this word?

An analogous case can be made with the other examples of (33) as well. In (33b), the transitive VN *koyoo* means 'employ', but the noun *koyoo-sya* 'employee' originally refers to a person who is employed rather than a person who employs someone, the latter meaning being represented by *koyoo-nusi* [employ-master], a hybrid word of S-J and native elements. Likewise, *saiyoo-sya* in (33c) (< *saiyoo* 'hire, adopt') means 'a hired person', *kaiko-sya* in (33d) (< *kaiko* 'dismiss, fire') 'a dismissed person', and *siyoo-nin* in (33e) (< *siyoo* 'use') 'an employee' or 'a servant'.

Examples of this sort appear to be limited to particular lexical items, and the "patient" interpretation of agent nominals is neither systematic nor productively available. On the contrary, the patient interpretation is allowed only under certain conditions, one of which is that both the agent and the patient that participate in the event denoted by the base predicate such as arresting and employment must denote humans, as required by the intrinsic meaning of *-sya* and other suffixes that denote humans by themselves. Moreover, they are mostly paired with lexical counterparts that clearly denote agents, as in *koyoo-sya* 'employee' and *koyoo-nusi* 'employer', or *siyoo-nin* 'employee' and *siyoo-sya* 'employer'.

In light of the fact that one and the same suffix may have an ambivalent function, as in *-sya*, which denotes an agent in *siyoo-sya* 'employer' but a patient in *koyoo-sya* 'employee', it would not be feasible to attribute the ambivalent interpretations to a structural property (such as argument structure) of the suffixes themselves. A semantic or pragmatic account would be more promising. For example, we could say that if the patienthood is clearly inferred from the lexical meaning of the base together with conventional social situations such as employment and layoff, the nominal may refer to a patient. A clear case is *taiho-sya* 'arrestee, one who is arrested'. The central participant in "arrest" situations is the person who gets arrested, since the agent is understood to be the police, so *taiho-sya* denotes an arrestee. *Yoogi-sya* [suspect-person] 'a suspected person' or 'a suspect' is a similar

case. In cases where the agent-patient relationship is explicit in the lexical meanings of bases, *-sya* refers to an agent or a patient according to the bases. In (34), for example, since the bases *kagai* 'inflict harm' and *higai* 'suffer harm' lexically express the distinction of agent and patient clearly, the derived nominals properly refer to the "doer" and "experiencer" involved in the harm.

(34) a.   *kagai-sya*   [inflict.harm-person]   'offender'

     b.   *higai-sya*   [suffer.harm-person]   'victim'

Our assumption that the agent-patient ambivalence is semantic or pragmatic in nature gains support from the recent practice (recommended by law) of using the prefix *hi-* 被 meaning 'undergo, suffer' to avoid ambiguity, as illustrated in (35).

(35) a.   *koyoo-sya* 'employer'   vs.   *hi-koyoo-sya* 'employee'

     b.   *siyoo-sya* 'employer'   vs.   *hi-siyoo-sya* 'employee'

The prefix *hi-* may even be attached to *taiho-sya* 'arrestee' and *yoogi-sya* 'suspect' to produce *hi-taiho-sya* 'arrestee' and *hi-yoogi-sya* 'suspect'.

Actually, however, agent-patient ambivalence is observed rather commonly in Japanese nominalization including compounds such as *hooti-zitensya* [abandon-bicycle] 'abandoned bicycle' or 'a bicycle which someone has abandoned on the street', and *taiho-reki* [arrest-record] 'criminal record' or 'a record that describes how many times a person was arrested in the past' (Kageyama 1996). The compound *taiho-keikan* [arrest-police.officer] may even mean 'an arrested police officer' rather than one who makes an arrest.

To recapitulate, Japanese agent suffixes uniquely refer to agents of action because they are grammaticalized from lexical nouns representing humans literally or figuratively. In this respect, they differ from the English suffix *-er*, which may refer to a variety of semantic relations other than agent. The only case where Japanese agent suffixes refer to non-agents is when they denote patients with certain VNs that select a human as their object. Because of this difference, the external argument restriction that has been proposed for the English *-er* nominals has no effect on Japanese agent nominals. The semantics of agent nominals in Japanese is determined by the lexical properties of individual suffixes.

4 Event vs. non-event distinction in agent nominals

As briefly mentioned in Section 1, the class of agent nominals discussed in this chapter differ from simple human nouns like *syatyoo* 'company president' and *titioya* 'father' in that they refer in one way or another to the actual involvement in

a particular action, event, or state represented by the base predicate. To see what is meant by "involvement in a particular action, event, or state", let us compare *unten-syu* [drive-hand] 'driver' and *unten-sya* [drive-person] 'driver'. As observed by Miyajima (1997), the former (*unten-syu*) designates a person who is engaged in the profession of driving, as it more often than not occurs as the head of compounds like *takusii-unten-syu* 'taxi driver', *basu-unten-syu* 'bus driver', and *okakae-unten-syu* 'hired driver, chauffeur'. It is the person's job to drive a certain car, bus, or truck, but he/she might not have actually driven the vehicle at all. For example, it perfectly makes sense to say that I hired the man as my *unten-syu* (chauffeur), but he quit without even touching the door of my car. The latter (*unten-sya*) is different in this respect. It does not represent the job or role as a driver but instead denotes a person who is actually involved in driving a vehicle at a particular place and at a particular time.

This different modality of action interpretation in agent nominals is actually a controversial issue that has been extensively discussed in the literature on English (Roeper 1987; Rappaport Hovav and Levin 1992; Alexiadou and Schäfer 2010). Many researchers agree that agent nominals are divided into two subclasses based on whether or not they refer to an actual event that is denoted by the base verb. The subclasses are dubbed "event nominal" and "non-event nominal". Event nominals denote persons who have actually been engaged in an action and thus presupposes that the event involving that action has occurred, while non-event nominals denote professions, roles, or instruments and do not presuppose that the event has actually occurred. The event/non-event distinction is also embedded in Pustejovsky's (1995) and Busa's (1996) analyses of nominalization in the Generative Lexicon framework. Pustejovsky proposes that the two classes of agent nominals can be subsumed under the well-known distinction in formal semantics between "stage-level" predicates (i.e. predicates that describe temporary or transitory events, as in *Firemen are available*) and "individual-level" predicates (i.e. predicates that depict more-or-less permanent or constant properties or attributes of entities, as in *Firemen are intelligent*: see Krifka et al. 1995). In Alexiadou and Schäfer (2010), the two classes of nominals are termed "episodic" and "dispositional", where episodic nominals correspond to event or stage-level nominals (i.e. those agent nominal that denote a person who conducts an action at a particular time) and dispositional nominals corresponds to non-event or individual-level nominals (i.e. those agent nominal that are interpreted as "someone intended or designed to V") in the same way that instruments and devices are designed for a specific purpose. As such, dispositional readings of agent nominals have the meaning of either specialized profession or specialized purpose. For the purpose of terminological consistency, I will henceforth use the terms "stage-level" and "individual-level" to characterize the two classes of agent nominals. Table 2 clarifies the terminological correspondences.

Table 2: Terminological clarification

	actual happening	**profession or role**
standard event semantics	stage-level	individual-level
Rappaport Hovav and Levin (1992)	eventive	non-eventive
Alexiadou and Schäfer (2010)	episodic	dispositional

What is particularly noteworthy about English *-er* agent nominals is the fact that they often permit ambiguity between a stage-level interpretation and an individual-level interpretation. For example, *driver* can be construed as either an individual-level or a stage-level noun. A person may be referred to as a driver if his or her job is operating a vehicle, or if he or she is/was actually engaged in the activity of driving. The former is an individual-level reading and the latter is a stage-level reading. In marked contrast to the English *-er*, which is in itself non-committal to the stage-level/individual-level distinction, most agentive suffixes in Japanese make the relevant distinction as part of their intrinsic lexical properties. Thus, Japanese has two counterparts to the English *driver*, namely *unten-sya* and *unten-syu*, where *-sya* and *-syu* attach to one and the same base *unten* 'drive' to give rise to a stage-level agent nominal in the former (*unten-sya*) and an individual-level agent nominal in the latter (*unten-syu*).

Given that the Japanese lexicon involves distinct lexical entries for stage/individual interpretations of agentive nominal, Kageyama (1999, 2002) observes that Japanese has a large set of suffixes that implement the semantic distinction in question. Table 3 gives a partial list of examples showing the difference in suffixes with the same bases.

Table 3: Stage-level and individual-level distinction by suffixes

Base predicate	**stage-level**	**individual-level**
VN *unten* 'drive'	*unten sya* 'driver'	*unten syu* '(pro) driver'
VN *ensoo* 'play (music)'	*ensoo sya* 'music player'	*ensoo ka* '(pro) musician'
VN *kango* 'nurse'	*kango nin* 'nurse'	*kango si* '(pro) nurse'
S-J morpheme *saku-* 'create'	*saku sya* 'author'	*sak ka* 'novelist'
infinitive *odori* 'dance'	*odori te* 'dancer'	*odori ko* '(pro) dancer'

The contrast of interpretation that we saw in the *-syu* and *-sya* suffixes above can also be found in Japanese base-ball terms. The players on offense, *batter* and *runner*, are called *da-sya* [hit-person] and *soo-sya* [run-person], respectively, where the agent nominals are formed with the suffix *-sya*, which, as in the case of *unten-sya* 'driver', expresses stage-level agent nominal that are performing the act at the time

of utterance. On the other hand, the players on defense, such as *pitcher*, *catcher*, and *infielder*, are translated into *too-syu* [pitch-person], *ho-syu* [catch-person], and *naiya-syu* [infield-person], respectively. The suffix *-syu* describes these terms as individual-level attributes or "roles" of the relevant players. In a baseball game, each player in the offence takes turns at bat and runs while players on defense play their designated positions. The difference in the function of players on offence and on defense is thus neatly characterized by the suffixes *-sya* and *-syu*.

This being said, however, it is not always the case that the semantic distinction described above coincides perfectly with the morphological distinctions. Japanese has plenty of noun-forming suffixes that make an individual/stage-level distinction, but those morphological markers do not always reflect the semantic distinction. The suffixes *-sya*, *-nin*, and *-nusi*, which basically derive stage-level agent nominals as in (36), are also used in deriving individual-level nominals as in (37).

(36) a. *sinsei-sya* [apply-person] 'applicant'

b. *hokoo-sya* [walk-person] 'pedestrian'

c. *tuukoo-nin* [pass-person] 'passer-by'

d. *kega-nin* [injure-person] 'injured person'

(37) a. *gizyutu-sya* [technic-person] 'technician'

b. *yuusiki-sya* [learning-person] 'learned person, expert'

c. *koosyoo-nin* [notarization-person] 'notary public'

d. *syoku-nin* [craft-person] 'craftsman'

The nominals in (36) describe persons involved in transitory events or states, whereas those in (37) depict roles or professions that will not easily change as time progresses. Importantly, the same suffixes (*-sya* and *-nin*) are split between the two sets of examples.

In the domestic tradition of Japanese grammar, Miyajima (1997) was the first to notice similar semantic differences in human nouns, making his original proposal that agent nominals in Japanese are classified in four categories in terms of their "aspectual" properties, as shown in (38).

(38) I. Actual Actor (*genjitsu-no shite*): 'X is/was actually doing something'
e.g. *kenbutu-nin* [sightsee-person] 'sightseer', *tuukoo-nin* [pass.by-person] 'passer-by'

II. Potential Actor (*senzaitekina shite*): 'X has potential to do something'
e.g. *unten-syu* [drive-person] 'driver', *sityoo-sya* [view-person] 'TV viewer', *syoosetu-ka* [novel-house] 'novelist'

III. Experiencer (*keiken-sha*): 'X has the experience of doing something'
e.g. *mokugeki-nin* [witness-person] 'eyewitness', *tyo-sya* [write-person] 'author', *hakken-sya* [discover-person] 'discoverer'

IV. Person in a certain state (*ittei-no jōtai-ni aru hito*)
e.g. *byoo-nin* [sick-person] 'sick person', *mai-go* [be.lost-child] 'lost child', *situgyoo-sya* [lose.job-person] 'unemployed person'

Miyajima's first category is Actual Actor (*genjitsu-no shite*), where the person is actually involved in the event denoted by the base of the nominal. The nominals in this category can be paraphrased as "X is or was doing something". The second category is Potential Actor (*senzaitekina shite*), where a person has potential to do an action denoted by the base. The third category is Experiencer, where a person is assumed to have experienced a situation or action denoted by the base of the nominal (i.e. "X has done/experienced something"). Miyajima's fourth category includes nominals that refer to a person who is in a certain state (i.e. "X is in a state of something"). Miyajima's work is insightful and highly suggestive in that it sheds fresh light on the relationship between the person denoted by an agent nominal and an event denoted by the nominal. Yet, there is room for discussion as to whether it is the "aspectual" properties that are responsible for the different categories as Miyajima claims.

Kageyama (2002) calls Miyajima's analysis into question on the grounds that his four-way classification of human nouns is merely an aspectual classification of the base of derived words. Alternatively, Kageyama claims that what is crucially relevant is the stage-level/individual-level distinction of nominals, as proposed by Pustejovsky (1995) and Busa (1996). He argues that Miyajima's Actual Actors and Potential Actors correspond to stage-level and individual-level nominals, respectively. He also notes that there is a third type of agent nominal, which he calls "event-occurrence nominals (*dekigoto hassei meishi*)". This type of nominal denotes an individual entity, yet it appears in syntactic constructions in parallel with event nouns. Thus, this type of nominal can co-occur with the verb of existence *iru* 'be' that selects animate subjects, as well as the existence verb *aru* 'be' that selects inanimate subjects. Observe the example in (39a) where both types of verb can be predicated of the event-occurrence nominal, in contrast with the example in (39b) where the stage-level nominal does not co-occur with the verb *aru*.

(39) a. *Hiruma wa ooku no zyokyaku ga iru/aru.*
day.time TOP many GEN passenger NOM be.ANM.PRS/be.INAN.PRS
'There are many passengers in the daytime.'

b. *Kuruma ni wa unten-syu ga iru/*aru.*
car LOC TOP drive-person NOM be.ANM.PRS/*be.INAN.PRS
'There is a driver in the car.'

Under Kageyama's proposal, agent nominals are divided into the following three types: individual-level nominals, stage-level nominals, and event-occurrence nominals. We will return to this discussion of event-occurrence nominals in Section 4.

Now, in their discussion on discussion on "event" and "non-event" agent nominals, Rappaport Hovav and Levin (1992) point out a striking difference that systematically holds for their two classes of nominal, which is that the agent nominals with event interpretation must take phrasal complements whereas those with non-event interpretation may not. This point is illustrated in (40) below, where the event interpretation correlates with the presence of complement structure.

(40) a. grinder of imported coffees

  b. a grinder

  c. a coffee grinder

In (40a), the argument of the base verb *to grind* is realized as the complement *of imported coffees*, which promotes the event (or stage-level) interpretation of the nominal *grinder*. On the other hand, the same nominal receives only a non-event (or individual-level) interpretation in (40b), where the argument is absent, and in (40c), where it is expressed word-internally in compound structure. This shows that a stage-level nominal must realize its argument (object) if there is one in order to have an event interpretation.

Such a correlation of the presence of event interpretation and the way arguments are realized in nominalization was noted by Sugioka (1986) in her monograph discussing Japanese agent nominals. Sugioka assumes that the difference in argument selection arises because agent nominals take over argument structure from the base in much the same way that action (event) nominals do (cf. Roeper 1987). She thus points out that the agent nominal *kaki-te* [write-person] 'writer' must express its object as a genitive NP overtly, as in (41a), or as a zero pronominal anaphorically understood in the discourse; otherwise, it is ruled out as shown in (41b).

(41) a. *Kare wa kono geki no kaki-te da.*
      he TOP this play GEN write-hand COP
      'He is the writer of this play.'

  b. **Kare wa kaki-te da.*
      he TOP write-hand COP
      'He is a writer.'

The deverbal compound *mono-kaki* [thing-write] 'writer' makes a sharp contrast in this regard, rejecting a phrasal complement altogether, as shown in (42b).

(42) a. *Kare   wa   mono-kaki   da.*
          he   TOP   thing-write   COP
          'He is a writer.'

   b. **Kare   wa   kono geki   no   mono-kaki   da.*
          he   TOP   this play   GEN   thing-write   COP
          'He is the writer of this play.'

Sugioka (1986) claims that the suffix *-te* bears the agent role of the base verb and implies a certain activity this agent engages in, while the compound *mono-kaki* type denotes a profession. The former takes over the verb's argument structure while the latter does not.

   Sugioka's observation holds generally for the distinction of stage-level vs. individual level nominals. In (43a), *sak-ka* 'writer, novelist', an individual-level noun, does not take a complement in syntactic structure, but it can take one in compound structure, as in (43b). On the other hand, in (44), the stage-level counterpart *saku-sya* 'author' takes a syntactic complement but fails to make up a compound with its complement noun, as observed by Kageyama (2002):

(43) a. **doowa   no   sak-ka*
          fairy.tale   GEN   write-house
          'a writer of fairy tales'

   b. *doowa-sak-ka*
          fairy.tale-write-house
          'a fairy-tale writer'

(44) a. *(sono)   doowa   no   saku-sya*
          (the)   fairy.tale GEN   write-person
          'the author of the fairy tale'

   b. **doowa-saku-sya*
          fairy.tale-write-person
          'a writer of fairy tales'

   Based on Sugioka's and Kageyama's analyses, it can be predicted that the stage-level nominal *unten-sya* 'driver' and its individual-level counterpart *unten-syu* 'driver' exhibit the same patterns as the *sak-ka/saku-sya* examples in (43) and (44). This prediction is partially borne out as in (45), where the compound form of the individual-level nominal is rejected.

(45) a. *torakku-unten-syu*   (individual-level)
　　　 truck-drive-hand
　　　 'a truck driver'

　　 b. **torakku-unten-sya*   (stage-level)
　　　 crashed.car-drive-person
　　　 lit. 'a crashed-car driver'

As shown in (46a), however, the individual level noun *unten-syu* can also take a genitive phrase *torakku no* 'of a truck', in parallel with the stage-level nominal *unten-sya* in (46b). If the genitive phrase in (46a) is really a complement of the head noun, then we should expect (46a) to be unacceptable.

(46) a. *torakku   no   unten-syu*   (individual-level)
　　　 truck   GEN   drive-hand
　　　 'a driver of a truck'

　　 b. *zikosya   no   unten-sya*   (stage-level)
　　　 crashed.car   GEN   drive-person
　　　 'a driver of a crashed car'

To account for such cases, Ono (2013) argues that there are two distinct modes of argument selection in nominals, depending on whether they are individual-level or stage-level. Individual-level nominals can select their arguments in a totally different manner than stage-level nominals. Their argument selection is based on lexical-semantic relations that are invoked by the denotations of the head noun.[5] Without going into the details of his argument, I will restrict my attention here to an eventive mode of argument selection by stage-level nominals.

5 Agent nominals and event nominals

In the preceding section, we saw that *-er* agent nominals, especially those that are classified as the stage-level type, exhibit certain parallelism with event or action nominalizations in that they allow modification of the event variable by a spatio-temporal adverbial and require the presence of a complement. As discussed by Grimshaw (1990), the complex event nominal *examination* in (47a) must take a

---

5 Along the same lines, the genitive phrases in *Nagoya no sityoo* 'the mayor of Nagoya', *ryokakki no pairotto* 'the pilot of airliner' are arguments that are selected via a different mode. Kageyama (2002: 50) notes *suirisyoosetu no sak ka* 'a writer of detective stories' is acceptable in contrast with *doowa no sak ka* 'a writer of fairy tales' in (46a), because *suirisyoosetu no* 'of detective stories' denotes a particular genre of stories that restrict the set of writers; hence, it is a restrictive modifier.

complement (*of the patients*), but the simple noun *exam* (or result nominal in Grimshaw's terminology) in (47b) cannot take a complement.

(47) a. The doctor's examination of the patients took a long time.

b. *The exam of the patients took a long time.

The contrast between (47a) and (47b) suggests that the capacity of taking complements is closely linked to the event interpretation of nominals. The head noun *exam* in (55b) totally lacks an implication of actual action, and so it does not allow a complement.

The mere parallelism between agent nominals and action nominalizations, however, does not mean that agent nominals can denote events, since, as Murphy (2011: 144) aptly notes, "The THING described by [the *-er* nominal] is more time-stable than the EVENT described by [the action nominal]". In this respect, Grimshaw (1990: 178, note 6) observes that *-er* agent nominals, unlike action nominalizations with the suffixes *-ing* and *-ation*, refer to individual entities, not events. The reason for this, she claims, is that although *-er* nominals can be modified by such frequency adjectives as *constant* and *frequent*, as in *a constant/frequent visitor*, they nonetheless reject locative, time, and instrumental adverbials. Compare thus the complex event nominal *induction* in (48a) with the *-er* nominal *inducer* in (48b).

(48) a. the induction of protein growth in a test tube/on Monday/with a technique

b. *the inducer of protein growth in a test tube/on Monday/with a technique

Baker and Vinokurova (2009) also note that a crucial difference between agent and action nominalizations is that the former rejects adverbial modification.

We have so far distinguished two types of agentive nominals, stage-level and individual-level nominals, in terms of event implications. Kageyama (2002) makes a further distinction, a third type of agent nominal, which he calls "event-occurrence nominals". As we saw in the previous section, event-occurrence nominals denote an individual entity, yet they behave like event nouns. Evidence in support of this claim comes from the existential constructions illustrated below (cf. Chapter 16 [Kishimoto, this volume] for a different analysis of these constructions). The agent nominals in (49a) can appear as the subject of the verb of existence *aru*, whereas those in (49b) cannot.

(49) a. {*sanka-sya/kesseki-sya/riyoo-sya/mokugeki-sya*}   *ga*   *aru*.
     participate-person/absent-person/use-person/witness-person NOM be.INAN.PRS
     'There are participants/absentees/users/witnesses.'

b. *{*unten-sya/hatumei-sya/kenkyuu-sya /ensoo-sya*}   *ga*.   *aru*.
     drive-person/invent-person/research-person/perform-person NOM be.INAN.PRS
     'There are drivers/inventors/researchers/performers.'

Event-occurrence nominals are not morphologically distinct from stage-level nominals, as shown in the above examples, but they differ in the event interpretation.

In order to understand how the nominals in (49a) are interpreted in the existential constructions, it is necessary to give a brief explanation of the verbs of existence. In Japanese, the verbs of existence have different forms depending on the animacy of the subject. *Iru* is used with an animate subject, and *aru* with an inanimate subject.

(50) a. *Nakaniwa   ni   gakuseitati   ga   iru/*aru.*
          patio   in   students   NOM   be.ANM.PRS/be.INAN.PRS
          'There are some students in the patio.'

    b. *Teeburu   no   ue   ni   hon   ga   aru/*iru.*
          table   GEN   top on   book   NOM   be.INAN.PRS/be.ANM.PRS
          'There is a book on the table.'

(50a) shows that *aru* cannot be used with the animate subject *gakuseitati* 'students'. Because of this animacy condition on the selection of verb form, the *aru* form is used when the subject is a so-called 'event noun'. In (51) the event nouns *kaigi* 'meeting' and *ziko* 'accident' co-occur with the verb *aru*.

(51) a. *San-zi   ni   kaigi   ga   aru.*
          three-o'clock at   meeting NOM   be.INAN.PRS
          'The meeting is at 3 o'clock.'

    b. *Kinoo   koosaten   de   ziko   ga   atta.*
          yesterday intersection at   accident NOM   be.INAN.PST
          'An accident occurred at the intersection yesterday.'

Keeping in mind the usage of the verb described above, let us now see what happens if an event-occurrence nominal occurs with *iru* and *aru*. Since the event-occurrence nominal in (52) is animate, it can co-occur with the *iru* form of the verb as in (52a). Thus, along with the *aru* verb in (52b), we have the *iru* form for event-occurrence nominal as illustrated below.

(52) a. *Tosyokan   ni   riyoo-sya   ga   ita.*
          library   in   use-person   NOM   be.ANM.PST
          'There are some visitors in the library.'

    b. *Tosyokan   ni   riyoo-sya   ga   atta.*
          library   in   use-person   NOM   be.INAN.PST
          'There was an occurrence of people visiting the library.'

The sentences in (52a) and (52b) are subtly different in meaning. (52a) simply conveys the meaning that some users are in the library, but (52b) means that an event takes place at the time of reference. Thus, with the verb *aru*, event-occurrence nominals are interpreted in a similar fashion as the event nouns in (51).

Given that event-occurrence nominals are distinguished from other agent nominals in terms of the selection of the existential verbs, let us see more evidence that event-occurrence nominals, though they denote individual agents, behave more like event nominals. First of all, we find that event occurrence nominals accept frequency adjectives as in (53b) in the same way that the event nominals in (53a) do. In sharp contrast to these are the stage-level nominals in (53c) and the individual-level agent nominals in (53d), which cannot accept frequency adjectives.

(53) a. *tamano*   {*hoomon/yasumi*}
         occasional   {visit/holiday}
         'an occasional visit/holiday'

     b. *tamano*   {*hoomon-sya/riyoo-sya*}
         occasional   {visit-person/use-person}
         'an occasional visitor/user'

     c. ***tamano*   {*unten-sya/siyoo-sya*[6]}
         occasional   {drive-person/use-person}
         'occasional {driver (in a stage-level sense)/user}'

     d. ***tamano*   {*siyoo-nin/unten-syu*}
         occasional   {use-person/drive-person}
         'occasional {employee/driver (in an individual-level sense)}'

In addition to the modification by frequency adjectives, event-occurrence nominals behave like event nouns in another respect. Recall that stage-level nominals can express their arguments as a complement to the head noun. Event-occurrence nominals can express adjuncts as well as complements from the verb they are derived from. The head noun *riyoo-sya* 'user' in (54) can have an argument 'the library' and an adjunct 'in wheelchairs', as we see below.

---

6 Note that *riyoo sya* 'user' in (53b) is an event occurrence nominal whereas *siyoo sya* 'user' in (53c) is regarded as a stage level nominal. However, as the English glosses show, they are sometimes synonymous. In particular, when *siyoo sya* is used in a context where it denotes a person who uses a public institution, as in *kooen no siyoo sya* [park GEN use person] 'user of the park', it is more likely to be interpreted as an event occurrence nominal. Thus, it is possible to use the verb *aru* with *siyoo sya* in the event occurrence interpretation.

(54) a. *tosyokan   no   riyoo-sya*
          library   GEN   use-person
          'the users of the library'

      b. *kurumaisu   de   no   tosyokan   no   riyoo-sya*
          wheelchairs   INS   GEN   library   GEN   use-person
          'the users of the library in wheelchairs'

Similarly, event nouns license both arguments and adjuncts. The noun *gappei-mondai* 'consolidation' in (55) is an argument and *kaigisitu de no* 'in the meeting room' is an adjunct.

(55) *kaigisitu   de   no   gappeimondai   no   giron*
      meeting room   LOC   GEN   consolidation   GEN   discussion
      'the discussion of consolidation in the meeting room'

Event-occurrence nominals and event nouns syntactically license adjuncts specifying for time, location, and means within the noun phrase. Thus, some agent nominals are more like event nouns than stage-level nominals. In other words, they are more "verb-like" than stage-level nominals. Kageyama (2002) claims that the parallelisms between event-occurrence nominals and event nouns are to be ascribed to the semantic representation of the two; more specifically, they have the Davidsonian event variable in their semantic representation, which licenses the spatio-temporal modification discussed above.

   Additional evidence in favor of the parallelism between event-occurrence nominals and event nouns is found in Ogawa (2001), where the differences among individual-level, stage-level, and event-occurrence nominals are discussed on the basis of event nouns, inalienable possession (IAP) nouns, and simple nouns. Some examples of each class of nouns are shown in (56).

(56) a. Event nouns
          tyuukoku 'advice', *teian* 'proposal', *kensa* 'examination'

      b. Inalienable possession nouns
          te 'hand', *kao* 'face', *kami* 'hair', *tume* 'nail', *asi* 'leg'

      c. Simple nouns
          kuruma 'car', *hon* 'book', *tukue* 'desk', *koppu* 'glass'

Based on an idea similar to the one proposed by Rappaport Hovav and Levin (1992), Ogawa argues that event nouns and IAP nouns contain an event position in their argument structure whereas simple nominals do not.

   Ogawa's proposal is partly motivated by what he calls the "multiple genitive construction". In (57), for example, the event noun *enzetu* 'speech' is modified by two genitive phrases *Oosaka-de no* 'in Osaka' and *John no* 'John's'.

(57) a. [**Oosaka de no** Zyon no enzetu] wa totemo omosirokat-ta.
　　　Osaka　LOC GEN John GEN speech TOP very　interesting-PST
　　　'John's speech in Osaka was very interesting.'

　　b. [Zyon no **Oosaka de no** enzetu] wa totemo omosirokat-ta.
　　　John GEN Osaka　LOC GEN speech TOP very　interesting-PST
　　　'John's speech in Osaka was very interesting.'

In (57), the spatio-temporal modifier 'in Osaka' is intended to modify the event noun *enzetu* 'speech' to mean that John delivered a speech in Osaka. What is interesting is that this interpretation obtains no matter where in the noun phrase the locative 'in Osaka' appears, i.e. on the left edge of the noun phrase as in (57a) or between the possessor ('John's') and the head ('speech') as in (57b). According to Ogawa, the flexible positioning of the spatio-temporal modifier is made possible by the event position (event argument) that the noun *enzetu* 'speech' contains in its argument structure.

　　In contrast, simple, non-event nouns like *inu* 'dog' do not exhibit this kind of flexibility in syntactic position. As shown in (58), the locative phrase is licensed only on the left edge of the NP; it is ruled out when it appears deep inside the NP, as in (58b), because the individual-level noun 'dog' does not have an event position.

(58) a. [**Kooen de no** Taroo no inu] wa totemo genki ga ii.
　　　park　LOC GEN Taro GEN dog TOP very　spirits NOM good
　　　'Taro's dog is in high spirits when he is out in the park.'

　　b. *[Taroo no **kooen de no** inu] wa totemo genki ga ii.
　　　Taro　GEN park LOC GEN dog TOP very　spirits NOM good
　　　'Taro's dog is in high spirits when he is out in the park.'

　　Assuming that Ogawa's distinction between event nouns and simple nouns essentially corresponds to our stage/individual distinction, we can take advantage of the contrast in (57) and (58) as a test frame to verify the validity of the two classes of agent nominals we have discussed so far. Consider the examples in (59) and (60), where *unten-sya* 'driver' (stage-level nominal) and *riyoo-sya* 'user' (event-occurrence nominal) take modifiers in two different positions.

(59) a. [**Gaikoku de no** sono saabisu no riyoo-sya]
　　　foreign.country LOC GEN the service GEN use-person

　　　wa kyuuzoo-si-teiru.
　　　TOP rapid.increase-do-ASP
　　　'The users of the service in foreign countries are increasing rapidly.'

　　b. [Sono saabisu no **gaikoku de no** riyoo-sya]
　　　the service GEN Foreign country LOC GEN use-person

　　　wa kyuuzoo-si-teiru.
　　　TOP rapid.increase-do-ASP
　　　'The users of the service in foreign countries are increasing rapidly.'

(60) a. [***Koosokudooro de   no***   *kuruma   no   unten-sya*]
          freeway   LOC   GEN   car   GEN   drive-person

      wa   inemuri   o   si-gati   da.
      TOP   sleep   ACC do-likely COP
      'Car drivers on the freeway are likely to fall asleep at the wheel.'

      b. *?[*Kuruma   no   **koosokudooro   de   no**   unten-sya*]
          car   GEN   freeway   LOC   GEN   drive-person

      wa   inemuri   o   sigati   da.
      TOP   sleep   ACC do-likely COP
      'Car drivers on the freeway are likely to fall asleep at the wheel.'

In (59b) and (60b), the locative modifiers intervene between the head and their complement phrase. The event-occurrence nominal in (59b) licenses the modifier in that position but the stage-level nominal in (60b) does not. If interchangeability of arguments and adjuncts serves as evidence for the presence of an event position in the nominal, as Ogawa claims, then (59) and (60) show that the former contains an event position, but the latter lacks it. This is completely on a par with Kageyama's (2002) claim that event-occurrence nominals contain an event variable in their semantic (event) structure.

6 Conclusion and future research perspectives

This chapter has shown the semantic and morphosyntactic properties of Japanese nominals in terms of three major theoretical issues arising from the previous discussions on English agent nominals: (i) the argument structure condition (the external argument generalization) on the interpretation of agent suffixes, (ii) the event and non-event distinction of agent nominals, and (iii) the parallelism between agent nominals and event nouns. The three issues can be put into questions addressed concerning agent nominals in Japanese. First, does the external argument generalization make sense in Japanese? If it does, unaccusative verbs would be rejected in the derivation of agent nominals. I have shown that this was not the case in Japanese. Japanese agent nominals can be derived from unaccusative, as well as unergative and transitive bases. Second, is there evidence to point to the distinction between eventive and non-eventive interpretations in agent nominals? I have shown that the event/non-event distinction which is sometimes morphologically marked in Japanese would offer significant theoretical and empirical insights into the nature of agent nominals. Finally, to what degree are agent nominals parallel with event nominals? In particular, though the exclusion of spatio-temporal adverbial modification is observed in agent nominalization (Baker and Vinokurova 2009), do Japanese

agent nominals behave in the same manner? We have seen that a subclass of eventive agent nominals exhibit a strong parallelism with event nouns in that they co-occur with a verb of existence that selects an inanimate subject and they are modified by adjuncts. Given this, it would serve as evidence against the claim that adverbial modification is generally excluded from agent nominals.

The biggest issue for future research, which may arise from the discussion in this chapter, would be the complexity of event implication in agent nominals. It is by now clear that a simple dichotomy of event vs. non-event nominals, which most of the previous research on this topic has suggested, is no longer an appropriate way to capture the eventiveness of agent nominals. This in turn suggests that an assumption based on the presence or absence of an event variable in event structure does not suffice to answer the question. Moreover, this issue is not solely concerned with agent nominals, but it ranges over other nominal types in general. For example, the distinction between individual-level/stage-level nominals would be valid for explaining the difference in the interpretation of non-human objects such as *wasure-mono* [lose-thing] 'lost property' and *nomi-mono* [drink-thing] 'a drink'. The former may be interpreted as a stage-level nominal while the latter, an individual-level nominal (you do not have to drink something before it is called 'a drink' but a lost article does not exist unless you have left something behind). Thus, the findings that we have uncovered regarding agent nominals would be more significant if we extend our analysis to other types of nominals as well.

Acknowledgments

References

Alexiadou, Artemis and Florian Schäfer. 2010. On the syntax of episodic vs. dispositional *-er* nominals. In Artemis Alexiadou and Monika Rathert (eds.), *The syntax of nominalizations across languages and frameworks*, 9–38. Berlin: Mouton de Gruyter.

Baker, Mark C. and Nadya Vinokurova. 2009. On agent nominalizations and why they are not like event nominalizations. *Language* 85. 517–556.

Bauer, Laurie. 1993. *English word formation*. Cambridge: Cambridge University Press.

Booj, Geert E. 1986. Form and meaning in morphology: The case of Dutch agent nouns. *Linguistics* 24. 503–517.

Busa, Federica. 1996. *Compositionality and the semantics of nominals*. Waltham, MA: Brandeis University dissertation.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Grimshaw, Jane and Armin Mester. 1988. Light verbs and θ-marking. *Linguistic Inquiry* 19. 205–232.

Ito, Takane and Yoko Sugioka 2002. *Gono shikumi to gokeisei* [The structure of words and word formation]. Tokyo: Kenkyusha.

Kageyama, Taro. 1982. Word formation in Japanese. *Lingua* 57. 215–258.

Kageyama, Taro. 1993. *Bunpō to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Syobo.

Kageyama, Taro. 1996. *Dōshi imiron* [Verb semantics]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 1999. *Keitairon to imi*. [Morphology and meaning]. Tokyo: Kurosio Publishers.

Kageyama, Taro. 2002. Dōsashumeishi ni okeru goi to tōgo no kyōkai [The boundary between lexicon and syntax in agentive nominals]. *Kokugogaku* [Japanese Linguistics] 53. 44–55.

Kageyama, Taro and Michiaki Saito. this volume. Chapter 1: Lexical strata and word formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Keyser, Samuel Jay and Thomas Roeper. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry* 15. 381–416.

Kobayashi, Hideki. 2004. *Gendainihongo no kangodōmeishi no kenkyū* [A study on Sino-Japanese verbal nouns in modern Japanese]. Tokyo: Hituzi Syobo.

Krifka, Manfred, Francis Jeffry Pelletier, Greg N. Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link. 1995. Genericity: an introduction. In Greg N. Carlson and Francis Jeffry Pelletier (eds.), *The generic book*, 1–124. Chicago: University of Chicago Press.

Kishimoto, Hideki. this volume. Chapter 16: Stative and existential/possessive verbs. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese Lexicon and Word Formation*. Berlin: De Gruyter Mouton.

Levin, Beth and Malka Rappaport. 1988. Nonevent *-er* nominals: A probe into argument structure. *Linguistics* 26. 1067–1083.

Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

Miyajima, Tatsuo. 1997. Hitomeishi no imi to asupekuto/tensu [The semantics of human nouns and aspect/tense]. In Yoshiaki Kawabata and Yoshio Nitta (eds.), *Nihongobunpō: Taikei to hōhō* [The Japanese grammar: System and methodology], 157–171. Tokyo: Hituzi Syobo.

Miyagawa, Shigeru. 1989. Light verbs and ergative hypothesis. *Linguistic Inquiry* 20. 659–668.

Miyamoto, Tadao and Hideki Kishimoto. this volume. Chapter 12: Light verb constructions with verbal nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Murphy, M. Lynne. 2011. *Lexical meaning*. Cambridge: Cambridge University Press.

Ogawa, Yoshiki. 2001. The stage/individual distinction and (in)alienable possession. *Language* 77. 1–25.

Ono, Naoyuki. 2013. Two modes of argument selection in nominals. In Bjarke Frellesvig and Peter Sells (eds.), *Japanese/Korean Linguistics* 20, 413–428. Stanford: CSLI Publications.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Ryder, Mary Ellen. 1999. Complex *-er* nominals: where grammaticalization and lexicalization meet. In Ellen Contini-Morava and Yishai Tobin (eds.), *Between grammar and lexicon*, 291–331. Amsterdam: John Benjamins.

Rappaport Hovav, Malka and Beth Levin. 1992. *-Er* nominals: Implications for the theory of argument structure. In Tim Stowell and Eric Wehrli (eds.), *Syntax and semantics 26: Syntax and the lexicon*, 127–153. New York: Academic Press.

Roeper, Thomas. 1987. Implicit arguments and the head-complement relations. *Linguistic Inquiry* 18. 267–310.

Shimamura, Reiko. 2006. Dōsashu o arawasu -er meishi ni tsuite [On *–er* nouns denoting agents]. In Hiroshi Yonekura (ed.), *Eigo no gokeisei: Tsūjiteki/kyōjiteki kenkyū no genjō to kadai* [English word formation: Problems in diachronic and synchronic studies], 368–407. Tokyo: Eichosha.

Sugimura, Hirohumi. 1986. *–Sha*, *-ka* [-sya, -ka]. *Nihongogaku* 5. 92–96.

Sugioka, Yoko. 1986. *Interaction of derivational morphology and syntax in Japanese and English*. New York: Garland.

Sugioka, Yoko. 1992. On the role of argument structure in nominalization. *Language and Communication* 10. 53–80. Keio University.

Tsujimura, Natsuko. 1990. Ergativity of nouns and case assignment. *Linguistic Inquiry* 21. 277–287.

Yumoto, Yoko. this volume. Chapter 9: Conversion and deverbal compound nouns. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Yuji Nishiyama
18 Complement-taking nouns

1 Introduction

Ordinarily, a noun can be used by itself, but there are also nouns that are semantically incomplete standing alone. For example, if one simply hears, "I saw a policeman there," the image of "a policeman" comes to mind. However, just hearing, "I saw a brother there" out of the blue, it is difficult to imagine what kind of brother is referred to and some sort of additional expression, such as *John's* in "I saw John's brother there" or *of my wife's grandfather* in "I saw a brother of my wife's grandfather there" is necessary. Let us refer to nouns like *brother* as "complement-taking nouns" and to the expressions that supplement the meaning of *brother*, like *John's* or *of my wife's grandfather* as the "complement" of *brother*. Japanese is rich in complement-taking nouns and there are situations in which they not only show assorted semantic relations but also affect syntactic uses as well. This phenomenon is especially interesting as a characteristic suggesting continuity between morphological and syntactic structure in Japanese. This chapter will outline the semantic and syntactic characteristics of this kind of noun and will examine the directions analyses of such nouns have taken and are taking.

It is first necessary to clarify why complement-taking nouns present a problem and what their linguistic significance is. Generally, verbs take arguments. For example, the transitive verb *devour* has the lexical semantic structure **eat.greedily** (e, x, y), where *e* is the so-called "Davidsonian event argument" and *x* and *y* are variables corresponding to the subject and the object, respectively. As a transitive verb, therefore, *devour* requires that a subject and object be explicitly represented in a sentence like *Bill devoured the meal*. With such a transitive verb, the subject and object are generally represented in the semantic structure. In contrast to this, *dine* has the lexical semantic structure **eat** (e, x, MEAL) in which, while the subject *x* is a variable, the object is filled by MEAL, a constant. Thus, *dine* behaves syntactically as an intransitive verb, not taking an object. Verbs like *dine* are rather exceptional. Whereas it is usual for a verb to take a variable in its lexical semantic structure and for that variable to be realized in syntactic structure as the object, a noun, in contrast, generally represents an autonomous semantic concept by itself and does not contain a variable in its lexical semantic structure. Contrary to this generalization, however, a complement-taking noun like *brother* takes some sort of variable in its lexical semantic structure and its concrete realization can be viewed as a complement, like *John's*.

Perhaps another example will illustrate the concept of complement to a noun. One could say of someone named 'Hanako', for example, *Hanako wa zyosei dearu* 'Hanako is a woman' or one could say *Hanako wa imooto dearu* 'Hanako is a younger sister.' But the nouns *zyosei* 'woman' and *imooto* 'younger sister' are fundamentally different in nature. There is no problem with stating that Hanako is a member of the set of *zyosei* 'women', but the statement that Hanako is a member of the set of *imooto* 'younger sisters' is less acceptable. This is because, unless younger sister is qualified by "whose" younger sister, the concept of "the set of younger sisters" is not meaningful. Hanako has the relation "*x*'s younger sister" to some person (*x*). In this sense, the term "younger sister" always calls for an argument "in relation to a person (*x*)" in its lexical semantic structure and, as in *Taroo no imooto* 'Taro's younger sister', this argument is generally realized in the form of the "whose" followed by the possessive case marker *no*, though, of course, there are also cases in which the argument is clear from the context and is not expressed linguistically. The noun *zyosei* 'woman', on the other hand, does not call for an argument "in relation to *x*" in its lexical semantic structure. Although it may be modified by an attributive noun phrase as in *kono mati no zyosei* 'a woman of this town', it means only "a woman who has some relation (R) to this town" and the nature of the relation (R) is determined pragmatically and is open to many interpretations depending on the context. It is clear from this that it is necessary to distinguish "sortal nouns" like *zyosei* 'woman', which can determine a set alone, from nouns like *imooto* 'younger sister', which show a relative relationship between multiple persons. Most nouns, like *ringo* 'apple', *haiyuu* 'actor', *zitensya* 'bicycle', *syuukyoo* 'religion', or *ame* 'rain' belong to the former type. Such nouns do not include variables in their lexical semantic structures and have a meaning that is complete in isolation with no need to establish a relationship with another element in order to semantically interpret them. In contrast to this type, such nouns as *haha* 'mother', *ozi* 'uncle', *zyoosi* 'boss', *senpai* 'senior', *gen'in* 'cause', *honba* 'place of origin', *kokyoo* 'hometown', or *hannin* 'perpetrator' belong to the latter type. Nouns like these include variables in their lexical semantic structures, do not have complete meanings in isolation, and require another element with which to establish a relationship in order to be interpreted semantically. Such nouns are "complement-taking nouns".

In the literature, nouns that either semantically or syntactically require some sort of complement have been discussed under a variety of names, including "nouns of inalienable possession", "relational nouns", "relative nouns", "one-place nouns", "unsaturated nouns", or "nouns of relativity".

(1) a. Nouns of Inalienable Possession (cf. Nichols 1988; Tsunoda 1995, 2009)
te 'hand', *kata* 'shoulder', *kami* 'hair', *hyoosi* 'cover sheet', *zyobun* 'foreword', *mokuzi* 'table of contents', *handoru* 'steering wheel', *bureeki* 'brake', *mado* 'window', *huta* 'lid', *totte* 'knob', *sode* 'sleeve', *eri* 'collar', *(kyoku no) kooda* 'the coda (to a piece of music)'

b. Relational Nouns (cf. Ono 2009)
   tuma 'wife', *zyoosi* 'boss', *enzin* 'engine', *gen* 'string (of an instrument)',
   bangumi 'program'

c. Relative Nouns (cf. Sugioka and Kageyama 2011)
   huta 'lid', *seimon* 'main gate', *sode* 'sleeve', *kabe* 'wall', *peezi* 'page',
   kata 'shoulder', *syuyaku* 'leading role', *honba* 'place of origin'

d. One-place Nouns (cf. Iori 2007)
   tyosyo 'authored work', *aizin* 'lover', *sakusya* 'writer', *atama* 'head',
   kao 'face', *kata* 'shoulder', *totte* 'knob', *huta* 'lid', *otooto* 'younger brother',
   tuma 'wife', *daitooryoo* 'president', *honba* 'place of origin', *taihan* 'majority',
   seito 'pupil', *sikin* 'fund'

e. Unsaturated Nouns (cf. Nishiyama 1990, 2003)
   syuyaku 'leading role', *yuusyoosya* 'victor', *senpai* 'senior', *sakusya* 'writer',
   sidookyoozyu 'academic advisor', *imooto* 'younger sister', *tuma* 'wife',
   hitudokusyo 'required reading', *gen'in* 'cause', *riyuu* 'reason', *honba* 'place
   of origin', *ketten* 'faults', *oturi* 'change', *kokyoo* 'home town', *teki* 'enemy',
   hannin 'perpetrator'

f. Nouns of Relativity (cf. Teramura 1992; Okutsu 1974)
   oturi 'change', *batu* 'punishment', *gen'in* 'cause', *kekka* 'result',
   tonari 'neighbor', *ue* 'top', *mae* 'front'

Many among these categories have not been given clear definitions and there are
many overlaps among the examples given for each. Section 2 first offers an overview
of the best known group, the nouns of inalienable possession, followed in Section 3
by an introduction to the groups conceptually closest to inalienable possession
nouns: relational, relative, and one-place nouns. Section 4 introduces unsaturated
nouns. Finally, Section 5 summarizes the discussion and examines remaining problems.

2 Nouns of inalienable possession

As an example of nouns that are semantically uninterpretable in isolation, this
section takes up nouns showing inalienable possession.

2.1 Body-part nouns and the whole-part relation

The Japanese genitive particle *no* can show a wide variety of relations between
two nouns in the pattern "NP1 *no* NP2", as seen in the such examples as *sensei no*

kuruma 'the teacher's car', *Tookyoo no tyuusinbu* 'the center of Tokyo', *kinoo no sinbun* 'yesterday's newspaper', or *zidoosya no yusyutu* 'automobile exports', but the archetypical relation expressed is that of NP1 as possessor and NP2 as possessum or possessed object, as shown in (2a) and (2b), below.

(2) a. *Taroo no {megane / ie / kuruma / pasokon}*
         Taro   GEN {eyeglasses / house / car   / PC}

     b. *Taroo no {te / asi / kao / kami}*
         Taro   GEN {hand / leg / face / hair}

In (2a) there is no necessary relation between the possessor and the possessed object; even if, say, *megane* 'eyeglasses' or *kuruma* 'car' were physically separated from Taro, they could still fulfill their original functions. This kind of temporary, coincidental possession is called "alienable possession". On the other hand, since *te* 'hand' or *asi* 'leg' in (2b) form part of the possessor Taro's body, they cannot fulfill their original functions if separated from the possessor. The link between the possessor (Taro) and the possessed (body-part nouns) in (2b) is necessary and it is normally impossible to sever this possessive relationship between possessor and possessed. This relation is called "inalienable possession", and body-part nouns are considered representative of "inalienable nouns" (see Haiman (1985: 130); Nichols (1988: 557–609, 1991: 119); Tsunoda (1995: 576); Dixon (2009: 262–312); Barker (2011: 1109–1130), among others).

   However, when considering body-part nouns, more important than inalienability is the whole-part relation. For example, *te* 'hand' constitutes a part of the "upper limbs" of a human (or human-like animal) and the upper limbs further constitute part of "the upper part of the body". Furthermore, "the upper part of the body" is part of the "body as a whole".

   The whole-part relation is also called "meronymy" (Cruse 1986: 157–180) and nouns like those in (3) also exemplify the whole-part relation.

(3) a. *nabe no {huta / totte / soko}*
         pan   GEN {lid   / knob / bottom}

     b. *kuruma   no {handoru   / zaseki / taiya}*
         automobile GEN {steering.wheel / seat   / tire}

     c. *heya no {mado / tenzyoo / kabe / doa}*
         room GEN {window / ceiling / wall / door}

     d. *sebiro no {sode / eri / poketto}*
         suit   GEN {sleeve / collar / pocket}

Each of these examples also shows the inalienable possession relation in that, if the NP2 noun were physically separated from NP1, it would become unable to fulfill its original function.

The whole-part relation in (2b) and (3) can be stated as "NP2 is a part of NP1", and in many cases there is another noun expressing another part that is involved in the tie between NP1 and NP2. For example, a *room* is a component part of a house (or other building) and one component of a *room* is a *door*, which further has parts such as a *doorknob*. In general, the knowledge of a set of hierarchical semantic relations expressing whole-part relations, as in HOUSE > ROOM > DOOR > DOORKNOB, is stored in the mind as part of the information in the lexicon. Because of this hierarchical relation, there exists a transitive relation such that, since a door is part of a room and a doorknob is part of a door, then a doorknob must also be part of a room. The same can be said of body-part nouns as well, so that a noun *hand* stands in the hierarchical relationship: BODY > UPPER PART OF BODY > UPPER LIMBS > HAND. As a result, since the *hand* is part of the *upper limbs* and the *upper limbs* are part of the *body*, the implication, "the hand is part of the body", can be drawn.

However, there is a difference in how body-part nouns are expressed as compared to the nouns in the examples in (3). The nouns in a hierarchical relation in (3) can be expressed with the nouns linked by *no*, as in *heya no doa no doanobu* 'the room's door's doorknob', but body-part nouns cannot be expressed in this manner, as shown by the unacceptability of **karada no zyoosi no te* 'the body's upper limbs' hand'.

As described above, body-part nouns and other nouns expressing parts are supported lexically by a set of hierarchical whole-part relations. Thanks to this sort of semantic underpinning, while nouns expressing parts and those expressing wholes are in an inalienable relation, as a matter of language use, no particular problem arises with using *te* 'hand' or *doa* 'door' alone.

2.2 Attributes and characteristics

Among the nouns treated as being in an inalienable relation, there are some examples that are, strictly speaking, difficult to characterize as being in a whole-part relation. Consider the examples in (4) (Nishikawa 2013: 66) and (5).

(4) a. *Hanako no koe*
    Hanako GEN voice

  b. *ki no kage*
    tree GEN shadow

  c. *Marukusu no haka*
    Marx GEN grave

(5) *Tyomusukii   no   tyosaku*
    Chomsky   GEN   works

In these examples, the NP2 can be considered to be a product or creation of NP1 and the existence of NP1 is essential for the existence of NP2. In this sense, such nouns can be considered to be in a kind of inalienable relation.

In examples like those in (6), where the NP2 is an attribute, like shape, of NP1, as well, the relation can be considered to be inalienable in the sense that the existence of NP1 is a prerequisite for that of NP2.

(6)   a.   *kodomo no   {sintyoo / taizyuu / taion}*
           child   GEN {height / weight / temperature}

      b.   *kuruma   no   {nagasa / seinoo   / iro}*
           automobile GEN {length   / performance / color}

      c.   *zyaketto no   {gara   / iro   / katati}*
           jacket   GEN {pattern / color / shape}

Sugioka and Kageyama (2011: 227) observe that nouns showing family relations as in (7) and those showing other human relations as in (8) have characteristics in common with inalienable nouns. For example, *haha* 'mother' does not typically simply indicate a woman, but is a concept that is first realized by the existence of a child who is in a direct relationship with the indicated woman. Similarly, the concept *tomodati (x)* 'friend (of *x*)' cannot exist by itself; it must designate as its reference point a person (*x*) who is not a blood relation but who is socially in a close relation with the person indicated.

(7) *Taroo no   {haha   / tuma / imooto   / ozi   / sobo   / senzo}*
    Taro   GEN {mother / wife / younger.sister / uncle / grandmother / ancestors}

(8) a. *Taroo no   {koibito / tomodati / tizin}*
       Taro   GEN {lover   / friend   / acquaintance}

    b. *Taroo no   {sisyoo / desi   / buka   / hisyo   / sidookyoozyu}*
       Taro   GEN {teacher / disciple / subordinate / secretary / academic.advisor}

Since there is no whole-part relation in examples like (4), (5), (6), (7), or (8), no transitive relation is established. That is, in *Taroo no tomodati₁ no tomodati₂* 'Taro's friend1's friend2', "friend2" cannot be said to be "Taro's friend".

It should be noted that examples like (9) are ambiguous between inalienable and alienable possession interpretations (see Chomsky (1972: 37–38) for a similar observation).

(9) *Taroo   no   te*
    Taro   GEN   hand
    'Taro's hand'

If (9) is interpreted as meaning 'Taro's hand as a part of his body', then this phrase shows inalienable possession, and this is the most natural interpretation of (9). However, other interpretations are possible depending on the context. As an example, suppose Taro is playing a game in which players who are blindfolded grab someone else's hand and then guess who has grabbed whose hand. In such a context, *Taroo no te* 'Taro's hand' in (10) could be interpreted as 'the hand that Taro has grabbed', and the relation between *Taroo* and *te* would be regarded as showing alienable possession (Nishikawa 2013: 75–76).

(10) *Taroo no   te   wa   Hanako no   te   da.*
     Taro   GEN hand TOP Hanako GEN hand COP.PRS
     'Taro's hand (the hand that Taro grabs) is Hanako's hand.'

However, this fact does not show that the word *te* 'hand' has two meanings or uses, one alienable and one inalienable. Lexically, *te* 'hand' has only the inalienable noun meaning because the *te* 'hand' in *Taroo no te* 'Taro's hand' in (10) is still, after all, part of someone's (e.g. Hanako's) body.

This section has presented Japanese examples of inalienable nouns, which can probably be found in all languages. It should be stressed that the concept of "inalienability" is fundamentally a matter of the word's lexical meaning and is not directly related to "whether or not the noun in question can be physically separated from its possessor". Even body parts like shoulders, hands, and feet, which are normally considered to be inseparable from the body, can in reality or in some possible world be physically separated (although, in such a case they would become unable to fulfill their original functions). The lid of a pot is frequently separated from the pot (but it then ceases to fulfill its original function of covering the pot). In short, every inalienable noun, such as *genkan* 'entrance hall, front door', must necessarily be specified in terms of its relation to its base, such as *ie* 'house' and, in this sense, the meaning of the noun *genkan* 'front door' is "inalienable" from the concept of *ie* 'house'. In that it is impossible to define the meaning of this sort of inalienable noun without reference to its base, they may be taken as nouns that require, in a broad sense, "variables" (or "arguments"). Of course, this does not preclude handling inalienability as a matter separate from the meaning it may have in individual languages, that is, as a matter of real or possible worlds or as a matter of cognitive motivation or of culture. On this point, refer to Tsunoda (1995: 576), Tsunoda (2009: 127), and Croft (1990: 175).

3 Noun groups close to inalienable possession nouns

This section will consider concepts close to that of inalienable possession nouns and examine their similarities and differences.

3.1 Relational nouns

When it is necessary to link a noun to another element in order to interpret the noun, it is called a relational noun. Examples include *syuyaku* 'leading role', *nedan* 'price', and *giseisya* 'victim'. In addition, family relational terms like *titi* 'father', *musuko* 'son', or *tuma* 'wife' can be said to be archetypical relational nouns.

In Section 2, it was stated that, in order to define inalienable possession nouns like *kata* 'shoulder' or *huta* 'lid', it was necessary to specify them in terms of their relation with their bases, the body or a container, and in this sense inalienable possession nouns can be considered to be a kind of relational noun.

A major question when considering relational nouns is how to interpret the requirement that they be linked to another element in order to receive an interpretation. Ono (2009: 258) defines relational nouns as in (11) and offers the examples in (12).

(11)   A relational noun is a noun whose extension is determined in relation to something else that exists separately.

(12)   a.   Nouns expressing family or social roles or relations (*tuma* 'wife', *zyoosi* 'boss')

   b.   Words expressing parts of the body (*atama* 'head' is part of someone's body)

   c.   Words expressing a component part of something (the extensions of *enzin* 'engine' or *gen* 'string' are determined by the relation of their being a part of some whole structure (*kuruma* 'automobile' or *gitaa* 'guitar').)

   d.   Words that express products (without a *seisakusya* 'producer', there would be no *bangumi* 'TV program')

Among Ono's examples in (12), there are some that do not fit the definition in (11). By "extension of a term A" in (11) is meant "the set of things to which it is correctly applied". For example, the extension of *inu* 'dog' is the set of all the dogs in the world. But the nouns expressing family or social roles or relations in (12a) cannot

alone determine an extension. That is, the extensions of *tuma* 'wife' and *zyoosi* 'boss' cannot be determined unless the values of *x* and *y* in "*x*'s wife" or "*y*'s boss" are set. Thus, these relational nouns satisfy the definition in (11). However, the extensions of the body-part nouns in (12b) can be determined by those nouns alone. For example, the extension of *mimi* 'ear', assuming knowledge of anatomy, can be determined without knowing whose ear it might be. Similarly, in order to determine the extensions of the nouns in (12c) one does not need to know which automobile the engine belongs to. In short, the differing sorts of characterizations given in (13) and (14) must be differentiated. Nouns like *mimi* 'ear' and *enzin* 'engine' that satisfy (13a) and (13b) will not necessarily satisfy (14).

(13) a. NP2 must be linked to NP1 in order to receive an interpretation.

b. NP2 first begins to function when it is a part of NP1.

c. The existence of the referent of NP2 presupposes the existence of NP1.

(14) The extension of NP2 is first determined by setting NP1.

In addition, it is not necessarily the case that nouns satisfying (13c) will always satisfy (14). The referent of *bangumi* 'program' will not exist unless its *seisakusya* 'producer' exists, but that does not mean that identification of the producer is necessary to determine the extension of *bangumi*. It is the same as saying that the referent of *isu* 'chair' will not exist unless its *seisakusya* 'producer' exists, but that does not mean that identifying the producer is necessary to determine the extension of *isu*. Furthermore, if the reason for considering *bangumi* 'program' a relational noun is that its producer must exist, then, as long as the existence of a producer is necessary for such things as *ronbun* 'dissertation', *tegami* 'letter', *syoosetu* 'novel', *nikki* 'diary', *kyoku* 'song', *pan* 'bread', *kuruma* 'automobile', and *ie* 'house' to exist, they must also be considered relational nouns. These nouns are not, however, normally considered as such. Considering this point, how to specify relational nouns appropriately is a matter for future research.

3.2 Relative nouns

Sugioka and Kageyama (2011: 225) distinguish between the two kinds of nouns discussed so far by the concepts of "autonomous" and "relative" nouns (it should be noted that the "relative nouns" referred to here are not the same as the "nouns that have relativity" of Teramura (1992)). "Autonomous" nouns are nouns like *neko* 'cat', *bin* 'jar', *ame* 'rain', or *yakan* 'teakettle' that designate autonomous concepts used alone. On the other hand, "relative" nouns are nouns like *huta* 'lid', *seimon* 'front gate', or *sode* 'sleeve' whose meanings are incomplete used alone and which must

be interpreted as having a relative relation to a concept with which it is linked. As stated in Section 2, *huta* 'lid', *seimon* 'front gate', or *sode* 'sleeve' are inalienable nouns in a whole-part relationship, but "relative" nouns include not only this sort of inalienable noun but also nouns showing family relations like *otto* 'husband', *tyoonan* 'eldest son', or *itoko* 'cousin', nouns showing other human relations like *tomodati* 'friend', *osananazimi* 'childhood friend', or *koibito* 'lover', and nouns denoting social roles like *sisyoo* 'teacher', *desi* 'disciple', *siyoosya* 'employer', or *siyoonin* 'servant' (Sugioka and Kageyama 2011: 227).

The standard for judging whether or not a noun is a relational noun is whether or not it can be used naturally in the frame *Kore wa X desu* 'This is an X' (Sugioka and Kageyama 2011). Taking, for example, an autonomous noun like *yakan* 'teakettle', it makes perfect sense used in the sentence *Kore wa yakan desu* 'This is a teakettle', but with a relative noun like *desi* 'disciple', if one were to suddenly be told *Kore wa desi desu* 'This is a disciple', one would want to respond *Dare no desi desu ka* 'Whose disciple?' Similarly, with a relative noun like *huta* 'lid', if someone were told *Kore wa huta desu* 'This is a lid', one would want to ask *Nan no huta desu ka* 'What is it a lid to?' This is because *huta* 'lid' is but one piece of a larger object that includes it. Accordingly, in order to fully comprehend what the noun is referring to, either the larger, whole object needs to be expressed with *no* as in *yakan no huta* 'the teakettle's lid' or it has to be possible to understand what the whole object is from the context. The same is true for other relative nouns (Sugioka and Kageyama 2011: 225). The "relative noun" concept is extremely close to the concept of "one-place" nouns, to which we now turn.

3.3 One-place nouns

Iori (2007:68) says there are nouns that require an argument marked by *no*, which he terms "one-place" nouns and those that do not require such an argument, which he terms "zero-place" nouns. *Tyosyo* 'authored work' and *aizin* 'lover' are examples of the former, while *hon* 'book' and *zyosei* 'woman' are examples of the latter. The distinction between "one-place" nouns and "zero-place" nouns overlaps conceptually to a large extent with the distinction between "relative" nouns and "autonomous" nouns (Section 3.2). According to Iori (2007: 68), the structural difference between one-place and zero-place nouns is linked to cohesiveness at the single-clause level. Consider the following examples (Iori 2007: 67).

(15) a. *Kinoo   sensei   wa   gakkai   no   kaizyoo   de*
       yesterday professor TOP   conference GEN   meeting.place LOC

   tyosyo   *ni   me   o   toosite   orare-ta.*
   authored.work   DAT   eyes ACC   pass.through   HON.be-PST
   'Yesterday the professor was looking over his books at the conference venue.'

b. *Kinoo   sensei   wa gakkai   no   kaizyoo   de*
yesterday professor TOP conference GEN meeting.place LOC

hon *ni   me   o   toosite   orare-ta.*
book DAT eyes ACC pass.through HON.be-PST
'Yesterday the professor was looking over the books at the conference venue.'

(16) a. *Giin   wa **aizin** to   hoteru kara dete.ki-ta*
council.member TOP lover with hotel   from emerge-PST

tokoro   o   toosatu-s-are-ta.
situation ACC candidly.photograph-do-PASS-PST
'The council member was snapped coming out of a hotel with his lover.'

b. *Giin   wa **zyosei** to   hoteru kara dete.ki-ta*
council.member TOP woman with hotel   from emerge-PST

tokoro   o   toosatu-s-are-ta.
situation ACC candidly.photograph-do-PASS-PST
'The council member was snapped coming out of a hotel with a woman.'

Tyosyo 'authored works' in (15a) can only refer to *sensei no tyosyo* 'the professor's works', but *hon* 'books' in (15b) is not interpreted as *sensei no hon* 'the professor's books', referring instead to some unspecified books. In the same way, *aizin* 'lover' in (16a) is interpreted as *giin no aizin* 'the councilor's lover', but *zyosei* 'woman' in (16b) cannot be interpreted as *giin no zyosei* 'the councilor's woman' but only as some unspecified woman. The difference between "one-place" nouns and "zero-place" nouns is also clear in the following examples.

(17) a. A: *Kinoo   mati   de   **sakka** o   mikake-ta yo.*
yesterday street LOC writer ACC see-PST   SFP
'Yesterday I saw a writer on the street.'

b. B: *Aa, soo desu ka.*
'Really? You don't say.'

(18) a. A: *Kinoo   mati   de   **sakusya** o   mikake-ta yo.*
yesterday street LOC author   ACC see-PST   SFP
'Yesterday I saw an author on the street.'

b. B: *#Aa, soo desu ka.*
'Really? You don't say.'

c. B:. *E? nan no? | #donna?*
oh what GEN / what.kind
'Oh? Of what? / #What kind?'

(19) a. A: *Kinoo　　hisasiburi ni **tyosyo** o　　yon-da　　yo.*
　　　　　yesterday a.while　　in opus　　ACC read-PST SFP
　　　　　'Yesterday I read (one of his) works for the first time in a long time.'

　　 b. B: *#Aa, soo desu ka.*
　　　　　'Really? You don't say.'

　　 c. B: *E?　dare　no?　/　#donna?*
　　　　　oh　who　GEN　/　#what.kind
　　　　　'Oh? Whose? / #What kind?'

If speaker A were to utter (17a) to B without any prior context, B could respond with (17b) and end the conversation. In contrast, if speaker A were to utter (18a) or (19a) to B without prior context, B cannot end the conversation with (18b) or (19b), respectively. Instead (18a) and (19a) lead to questions like those in (18c) and (19c). This is similar to how, in response to (20a), (20b) would not be an appropriate end to the exchange; instead something like (20c) is called for.

(20) a. A: *Kinoo　　yonda yo.*
　　　　　yesterday read　　SFP
　　　　　'Yesterday I read.'

　　 b. B: *#Aa, soo desu ka.*
　　　　　'Really? You don't say.'

　　 c. B: *E?　nani　o?*
　　　　　oh　what　ACC
　　　　　'Oh? What?'

Iori (2007: 150–157) argues that since, just as the verb *yomu* 'read' in (20) calls for a syntactic argument *nani* 'what', one-place nouns like *sakusya* 'author' (18a) and *tyosyo* 'authored work' (19a) call for the arguments *nan no* 'of what' and *dare no* 'by whom', respectively, the arguments called for by one-place nouns should also be considered "syntactic arguments". Furthermore, as seen in (21), unlike zero-place nouns, one-place nouns can also exert an inter-sentential cohesiveness.

(21) *Siba　Ryootaroo si　ga　　nakunat-ta.　　Nagai aida　{tyosyo/?hon}　ni*
　　　Shiba Ryotaro　Mr. NOM　pass.away-PST long　period {works/books} DAT
　　　sitasinde.ki-ta　watasi wa　hidoi syokku o　　uke-ta
　　　be.raise.on-PST I　　TOP awful shock　ACC receive-PST
　　　'Mr. Ryotaro Shiba died. Raised on (his) {works/?books}, as I was, it was an awful shock.'

Iori offered examples, like the following, of one-place nouns:

(22) (parts) *atama* 'head', *kao* 'face'; *huta* 'lid', *totte* 'knob'; (family) *otooto* 'younger brother', *tuma* 'wife'; (job titles) *daitooryoo* 'president', *butyoo* 'section chief'; (nouns of relativity) *ue* 'top', *mae* 'front'; (places) *kaizyoo* 'venue', *honba* 'place of origin'; (verbal nouns) *zitugen* 'realization', *syuuri* 'repair', *hakai* 'destruction', *kenkyuu* 'research'; (portions) *taihan* 'greater part', *zyuppaasento* '10%'; (other) *seito* 'pupil', *sakusya* 'author', *sikin* 'fund'

One-place nouns thus overlap heavily with inalienable possession nouns, relational nouns, and relative nouns. However, because the idea of one-place nouns is not a semantic notion but is a lexical and syntactic concept proposed to explain textual cohesion and because one-place nouns include verbal nouns like *zitugen* 'realization', *syuuri* 'repair', and *kenkyuu* 'research', they are conceptually different from the purely semantic notions of inalienable possession nouns, relational nouns, and relative nouns.

4 Unsaturated nouns

The previous sections have examined inalienable possession nouns, relational nouns, relative nouns, and one-place nouns as complement-taking nouns. However, there is another class of complement-taking nouns, called "unsaturated nouns", that show atypical behavior semantically and syntactically. In this section we will explain how "unsaturated nouns" are atypical even among complement-taking nouns.

4.1 What are "unsaturated nouns"?

Nishiyama (1990, 2003) argues the importance in semantics of the distinction between saturated and unsaturated nouns. As is well known, the extension of the adjective *akai* 'red' is the set of things to which *akai* 'red' is properly applied, that is, the set of things that are red. However, as the word *tosi-ue* 'older' includes in its meaning a variable X, as in *X yori tosi-ue* 'older than X', the set of things to which *tosi-ue* is properly applied cannot be determined without filling in a value for the variable X. That is, one cannot establish a "set of things that are older" unless the variable X is specified. In this sense, the meaning of *akai* 'red' is complete but the meaning of *tosi-ue* 'older' is not. The same kind of observation can also be made of nouns. Consider the examples in (23).

(23) a. Saturated Nouns: *gaka* 'painter', *sakka* '(professional) writer, novelist', *haiyuu* 'actor', *otoko* 'man', *inu* 'dog', *kubikazari* 'necklace', *ringo* 'apple', *zitensya* 'bicycle', *syuukyoo* 'religion'

    b. Unsaturated Nouns: *(X no) zyoosi* '(X's) boss', *(X no) sakusya* 'the author (of X)', *(X no) syuyaku* 'the leading role (of X)', *(X no) imooto* '(X's) younger sister', *(X no) syumi* '(X's) hobby', *(X no) kokyoo* '(X's) hometown'

The nouns in (23a) are complete by themselves and the extensions of these nouns, that is, the set of things to which each is properly applied, can be determined autonomously. Accordingly, the set of *gaka* 'painters' can be established and it is possible to speak of *gonin no gaka* 'five painters'. The truth or falsity of the statement *tonari no heya ni go-nin no gaka ga iru* 'there are five painters in the next room' can be determined with no previous context. Nishiyama (2003) calls this sort of noun "saturated" nouns.

However, the nouns of (23b) are different. For example, since the meaning expressed by the word *zyoosi* 'boss' includes a variable *X*, as in *X no (zyoosi)* 'X's (boss)', the extension of *zyoosi* 'boss' cannot be determined without filling in a value for the variable *X*. That is, as long as the *dare no* 'whose' of *dare no zyoosi* 'whose boss' is left unspecified, it is impossible to determine whether or not a particular person belongs to the set of persons to whom *zoosi* 'boss' is properly applied. Furthermore, in a situation in which the *dare no* 'whose' of *dare no zyoosi* 'whose boss' is unspecified, the truth or falsity of *tonari no heya ni go-nin no zyoosi ga iru* 'there are five bosses in the next room' cannot be determined. Accordingly, unless the value of *X* is understood from context, a "set of bosses" cannot be defined. The same is true of the other nouns in (23b). Since the extensions of such nouns cannot be determined autonomously as they include a variable X as part of their meaning, Nishiyama called this type of noun "unsaturated" nouns. In a nutshell, "saturated" nouns such as *gaka* 'painter', *ringo* 'apple', and *kuruma* 'car' have sets of individuals as their extensions, whereas "unsaturated" nouns such as *imooto* 'younger sister', *senpai* 'senior', *zyoosi* 'boss', and *kokyoo* 'hometown' are not able to have sets or individuals as their extensions.

The distinction between saturated and unsaturated nouns resembles the distinction between transitive and intransitive verbs in English. Saturated nouns are nouns that do not include a variable and, looking at English verb examples, can probably be considered to correspond to weather verbs like *rain* or *snow* which, although taking a subject *it* grammatically, do not require a variable semantically in order to express a complete statement. In the case of verbs, examples like *rain* and *snow* are rather unusual, but in the case of nouns, the situation of having all semantic arguments saturated and having no variables is the general case. On the other hand, unsaturated nouns are nouns having variables and, as with the verb *devour* given as an example in Section 1, they refer to ones for which argument-taking is obligatory. Put another way, unsaturated nouns can be said to be nouns that obligatorily

require a complement. Since Nishiyama (2003) termed the variable *X* that unsaturated nouns include in their semantic structure a "parameter", the same term will be used below as well. The term "parameter" used here refers to the variable required in the lexical semantic structure of an unsaturated noun, and it is with the provision of a value for this variable that the noun first comes to have an extension.

Since a large portion of the noun lexicon in Japanese is made up of saturated nouns, we will present here additional examples of unsaturated nouns in order to clarify their understanding. As shown in Table 1, unsaturated nouns are highly varied, including not only those expressing human relations, but also those showing time and space and even concrete and abstract concepts.

Table 1: Examples of unsaturated nouns

Human nouns	*imooto* 'younger sister', *ozi* 'uncle', *koibito* 'lover', *syuyaku* 'leading role', *sakusya* 'author', *sikaisya* 'master of ceremonies', *hannin* 'perpetrator', *senpai* 'senior', *zyoosi* 'boss', *syatyoo* 'company president', *sidookyoozyu* 'academic advisor'
Place nouns	*ue* 'top', *migi* 'right', *tonari* 'next door', *usiro* 'back', *minami* 'south', *nisi* 'west', *kaizyoo* 'venue', *kokyoo* 'hometown', *honba* 'place of origin, home'
Time nouns	*mae* 'before', *ato* 'after', *saki* 'earlier'
Concrete nouns	*teki* 'enemy', *hitudokusyo* 'required reading', *koobutu* 'favorite food', *senmon* 'speciality'
Abstract concept nouns	*riyuu* 'reason', *gen'in* 'cause', *mokuteki* 'goal', *batu* 'punishment', *kekka* 'result', *syooko* 'proof', *kuse* 'habit', *ketten* 'deficiency', *tatari* 'curse'

The distinction between saturated and unsaturated nouns is also reflected in the semantic relation between NP1 and NP2 in the "NP1 *no* NP2" pattern (Nishiyama 2003).

(24) a. *Hokkaidoo-syussin no   sakka* (*no* is a copula meaning "is")
     Hokkaido-native   COP   writer
     'a writer who is a native of Hokkaido'

  b. *Genzi Monogatari no   sakusya* (*no* is a genitive marker)
     Tale.of.Genji   GEN   author
     'the author of The Tale of Genji'

The *no* in (24a) is a copula meaning 'is' and the modifier *Hokkaidoo-syussin no* 'is a native of Hokkaido' has the function of selecting a subset from the head noun set of *sakka* 'writers'. However, the *no* in (24b) cannot be replaced by the copula and the phrase *Genzi Monogatari no* does not function to select a subset of the set of *sakusya* 'authors'. As discussed earlier, while a set of writers can become a topic for con-

sideration, with an unsaturated noun like *sakusya* 'author', no such set can be made into a topic. The modifier *Genzi Monogatari no*, by setting a value for the parameter of *sakusya* 'author', functions to create the whole completed meaning of *Genzi Monogatari o kaita hito* 'the person who wrote Genji Monogatari'.

The difference in the semantic relation between NP1 and NP2 in (25) also reflects the difference between the saturated noun *hon* 'book' and the unsaturated noun *hitudokusyo* 'required reading'.

(25) a. *gengogaku no hon*
      linguistics GEN book
      'a linguistics book'

    b. *gengogaku no hitudokusyo*
       linguistics GEN required.reading
       'required reading for linguistics'

In (25a) *hon* 'book' alone can define a set and *gengogaku no* 'linguistics' has the function of restricting reference to a subset within the set of "books". On the other hand, in (25b) because *hitudokusyo* 'required reading' is an unsaturated noun including a parameter like "required reading for *X*", it cannot define a set on its own and *gengogaku no* has the function of filling in a value for that parameter. Only by forming a saturated noun phrase like *gengogaku no hitudokusyo* 'required reading for linguistics' can a set be defined. The noun phrases "NP1 *no* NP2" in (26) have the relation [parameter value + *no* + unsaturated noun] as was the case in (24b) and (25b).

(26) a. *kono sibai no syuyaku*
      this play GEN leading.role
      'the leading role of this play'

    b. *ano kaisya no syain*
       that company GEN company.staff
       'the staff of the company'

    c. *Tokugawa-ke no teki*
       Tokugawa-clan GEN enemy
       'the enemy to the Tokugawa clan'

    d. *titi no koobutu*
       father GEN favorite.food
       'my father's favorite food'

    e. *Hanako no kokyoo*
       Hanako GEN hometown
       'Hanako's hometown'

f. *kaki-ryoori   no   honba*
   oyster-cuisine GEN   place.of.origin
   'the home of oyster cuisine'

g. *Taroo   no   imooto*
   Taro   GEN   younger.sister
   'Taro's younger sister'

4.2 Tests for unsaturated nouns

The following three simple tests can distinguish between saturated and unsaturated nouns. First, if it is possible to give an appropriate answer to a question of the form of (27) asked with no prior context, then the N is a saturated noun; if it is not possible to give an appropriate answer, then the N is an unsaturated noun (Yamaizumi 2013: 13–14).

(27) *{Anata / Are / Ano hito   / Kore} wa   N desu ka?*
      {you   / that / that person / this} TOP N COP Q
      'Is/Are {you / that / that person / this} an N?'

If the N is a saturated noun, then, since it does not contain a parameter in its lexical meaning, the attributes expressed by N can be determined by the noun alone. It is, therefore, possible to answer "yes" if the referent of the subject noun phrase has those attributes or "no" if it does not. For example, suppose a woman were asked the following.

(28) *Anata wa   kangosi desu ka?*
      you   TOP nurse   COP   Q
      'Are you a nurse?'

In response to this question, if the woman has a nursing certificate, she could answer "Yes" and if she does not, then "No". This fact shows that *kangosi* 'nurse' is a saturated noun.

     If, on the other hand, N is an unsaturated noun, then, since it includes a parameter in its lexical meaning, it is impossible to determine its attributes without filling in a value for the parameter. Accordingly, it is not possible to answer a question like (27) with no prior context. Suppose, for example, a woman were asked the following:

(29) *Anata wa   ane   desu ka?*
      you   TOP elder.sister COP   Q
      'Are you an elder sister?'

Without filling in a value for *dare* 'who', as in *dare no ane* 'whose elder sister', the question is unanswerable. This fact shows that *ane* 'elder sister' is an unsaturated noun. In English as well, there is no problem at all with (30a), but without some understanding of whose elder sister, (30b) is an unacceptable question.

(30) a. Are you his elder sister?

b. ?Are you an elder sister?

The following pairs of sentences make it clear that *haiyuu* 'actor', *sakka* 'writer', and *mori* 'forest' are saturated nouns while *syuyaku* 'leading role', *sakusya* 'author', and *honba* 'place of origin' are unsaturated. The # mark on the (b) sentences indicates that they are unacceptable without some special context.

(31) a. *Ano hito   wa   sakka desu ka?*
    that person TOP writer COP Q
    'Is (s)he a writer?'

  b. *#Ano hito   wa   sakusya desu ka?*
    that person TOP author   COP Q
    'Is (s)he the author?'

(32) a. *Anata wa   haiyuu desu ka?*
    you   TOP actor   COP Q
    'Are you an actor?'

  b. *#Anata wa   syuyaku   desu ka?*
    you   TOP leading.role COP Q
    'Are you the lead?'

(33) a. *Koko wa   mori   desu ka?*
    here TOP forest COP Q
    'Is this a forest?'

  b. *#Koko wa   honba   desu ka?*
    here TOP place.of.origin COP Q
    'Is this the place of origin?'

The explanation for (31), for instance, is as follows. The meaning of *sakka* 'writer' is roughly "the occupation of creating works like novels or poems (or a person engaged in this occupation)" and concerning "that person" it is entirely possible to question whether or not he or she is a writer without reference to any works he or she may have produced. That is, since (31a) can be uttered alone out of the blue,

sakka 'writer' is a saturated noun. With regard to this point, *sakusya* 'author' is fundamentally different from *sakka* 'writer'. Without specifying what work (such as a novel or poem) he or she may be author of, it is impossible to determine the truth or falsity of whether or not the person in question is the author. Therefore, unless some artistic work is already under discussion and the question is whether or not the person is the work's creator, the question in (31b) makes no sense. That is, the noun *sakusya* 'author' is an unfilled unsaturated noun and cannot be used alone out of the blue but must have a concrete work filled in as the value for the parameter *X no* 'of X'.

By this test, most of the inalienable possession nouns, relational nouns, and relative nouns examined so far are also unsaturated nouns. However, it also shows that whole-part relation nouns like body-part nouns are saturated nouns. Consider (34).

(34) *Kore wa   kaminoke desu ka?*
      this   TOP   hair   COP   Q
      'Is this hair?'

It is not the case that (34) is unanswerable without knowing whose hair is being asked about. Even without designating the possessor of the hair, if the object denoted by *kore* 'this' is hair, it is possible to answer "yes" and if it is not, to answer "no". This is completely unrelated to the fact that we can easily answer (34) because we have experiential knowledge of the archetypical prototype of "hair". Depending on the circumstances, there could be cases where it is difficult to determine whether or not some substance is hair or not and forensic tests are required. What is in question here is not whether or not it is easy to determine whether some substance is hair, but whether or not it is necessary to determine the possessor of the hair in order to judge whether the substance is or is not hair. Essentially, it is a question of whether or not *kaminoke* 'hair' is regarded as a word taking a parameter like *X no kaminoke* 'X's hair' and the substance is or is not hair depending on the value of the parameter. This is definitely not the case. Therefore, by this test, *kaminoke* 'hair' must be a saturated noun. Other body-part nouns like *te* 'hand' or *kata* 'shoulder' and expressions showing structural parts of something like *hyoosi* 'cover sheet', *handoru* 'steering wheel', *mado* 'window', *doa* 'door', or *yane* 'roof' are similar. Generally speaking, inalienable possession nouns showing whole-part relations are determined to be saturated nouns by this test.

A second test to determine whether a given noun is a saturated or unsaturated noun is to use the schema in (35) (Nishikawa 2013: 67–68).

(35) *Kono N wa {dare no   / nan no}   N desu ka?*
      this   N TOP {who GEN / what GEN} N COP Q
      'Whose N is this N? / What N does this N go with?'

If the same N is to be filled in for both the N in (35), it is possible to fill in saturated nouns like *kutu* 'shoes', *kookoku* 'advertisement', or *oto* 'sound' as in (36).

(36) a. *Kono ookina kutu   wa   dare no   kutu   desu ka?*
            these big   shoes TOP who GEN shoes COP Q
            'Whose shoes are these big shoes?'

   b. *Ano kawatta kookoku   wa   nan   no   kookoku   desu ka?*
         that weird   advertisement TOP what GEN advertisement COP Q
         'What is that weird advertisement an advertisement for?'

   c. *Ano oto   wa   nan   no   oto   desu ka?*
         that sound TOP what GEN sound COP Q
         'What was that sound?'

In each of the sentences in (36), *kono kutu* 'these shoes', *ano kookoku* 'that advertisement', and *ano oto* 'that sound' can refer to some object and the sentence asks for more information about that object. This is possible because the nouns filling in the N are saturated nouns and is impossible if N is an unsaturated noun, as shown below.

(37) a. ?*Ano zyoosi wa   dare no   zyoosi desu ka?*
            that boss   TOP who GEN boss   COP Q
            'Whose boss is that boss?'

   b. ?*Kono honba   wa   nan   no   honba   desu ka?*
         this   place.of.origin TOP what GEN place.of.origin COP Q
         'What is this place of origin the place of origin of?'

   c. ?*Kono ketten   wa   dare no   ketten   desu ka?*
         this   deficiency TOP who GEN deficiency COP Q
         'Whose deficiency is this deficiency?'

   d. ?*Ano gen'in wa   nan   no   gen'in desu ka?*
         that cause TOP what GEN cause COP Q
         'What is that cause the cause of?'

   e. ?*Kono teki   wa   dare no   teki   desu ka?*
         this   enemy TOP who GEN enemy COP Q
         'Whose enemy is this enemy?'

   f. ?*Ano kokyoo   wa   dare no   kokyoo   desu ka?*
         that hometown TOP who GEN hometown COP Q
         'Whose hometown is this hometown?'

Out of the blue, with no prior context, *ano zyoosi* 'that boss' is strange in and of itself. This expression would only be acceptable if the value X of the parameter *X no zyoosi* 'X's boss' were understood from context. But the predicate noun phrase of (37a) includes the interrogative *dare no* 'whose' and asks for that parameter value. The subject and predicate noun phrases in (37a) are inconsistent with regard to the value of the variable for *zyoosi* 'boss' and the sentence is therefore unacceptable. The same can be said of the other examples in (37).

Consider now the schema of (35) with *asi* 'feet', *huta* 'lid', and *yane* 'roof' filled in for N.

(38) a. *Kono asi   wa   dare no   asi   desu ka?*
         this   feet TOP who GEN   foot COP   Q
         'Whose feet are these feet?'

  b. *Kono huta wa   dono   nabe no   huta desu ka?*
      this   lid   TOP   which pan GEN   lid   COP   Q
      'Which pan is this lid the lid for?'

  c. *Kono yane wa   dono   ie   no   yane desu ka?*
      this   roof TOP   which house GEN   roof COP Q
      'Which house is this roof the roof of?'

(38a) might be said in a situation where someone's feet are sticking out from the bedcovers but the speaker does not know whose feet they are. (38b) is a sentence that might be uttered when one knows that this is a lid for some pan but does not know which pan it goes with. (38c) would be a natural sentence to ask when looking at an aerial picture of a town. The fact that the sentences of (38) are acceptable shows that it is not necessary to specify their possessors in order to interpret *kono asi* 'these feet', *kono huta* 'this lid', and *kono yane* 'this roof'. By this test as well, *asi* 'feet', *huta* 'lid', and *yane* 'roof' are saturated nouns. This fact does not, of course, conflict with the fact that, as inalienable possession nouns, *asi* 'feet', *huta* 'lid', and *yane* 'roof' cannot function separate from their possessors. The foregoing observation shows that the part-base (whole) relation seen in *asi* 'feet', *huta* 'lid', and *yane* 'roof' is different from the parameter-unsaturated noun relation found with nouns like *imooto* 'younger sister', *teki* 'enemy', and *honba* 'place of origin'.

The third test for whether a given noun is saturated or unsaturated is its ability to co-occur with numerical quantifiers. In the case of saturated nouns, it is naturally possible to attach numerical quantifiers in front of them, as in (39).

(39) a. *go-mei   no   kaisyain*
         five-CLF GEN   office.workers

  b. *san-dai   no   kuruma*
      three-CLF GEN   cars

There is no need to specify a company in order to say whether or not some person is an office worker. Accordingly, if there should be two office workers from company A and three office workers from company B, we can refer to them collectively as five office workers. The same is true of (39b). However, as stated in Section 4.1, in the case of unsaturated nouns, statements like those in (40) are impossible unless the parameter is set to a single value (Yamaizumi 2013: 13).

(40) a. *go-mei   no   syain*
          five-CLF GEN member.of.company

     b. *mittu no   riyuu*
          three GEN reasons

     c. *yottu no   teki*
          four   GEN enemies

When we interpret the sentences of (40) we tacitly fix the value of the parameters to the same values. For example, while (40a) may have the meaning "five members of a particular firm", it cannot mean "two members of firm A and three members of firm B". This is because *A-sya no syain* 'members of firm A' and *B-sya no syain* 'members of firm B' have different extensions. Similarly, (40b) could mean "three reasons Hanako got divorced" but it cannot be used to refer collectively to "one reason that Hanako got divorced" and "two reasons that Taro resigned from the company". The same is true of the *yottu no teki* 'four enemies' of (40c), which cannot be used by a country that is in the position of having to negotiate with two enemies of country A, one enemy of country B, and one enemy of country C to state its situation as "we have four enemies we have to negotiate with". This is because, essentially, numerical quantifiers can only modify expressions with parameters fixed to the same values. Consider now the following expressions.

(41) a. *mittu no   kao*
          three GEN faces

     b. *san-ko   no   huta*
          three-CLF GEN lids

     c. *rok-ko   no   mado*
          six-CLF GEN windows

(42) *Asoko ni   mittu no   kao   ga   miemasu   ne.*
      there   LOC three GEN faces NOM are.visible SFP

     Anata ga   mioboe   no   aru   kao wa   dore   desu ka?
      you   NOM recognition GEN exist face TOP which COP   Q
      'There are three faces visible over there, right? Which face looks familiar to you?'

Since *kao* 'face' is an inalienable possession noun and must be "someone's face", it may seem to be a noun requiring a parameter *X*. However, as shown by (42) it is possible to refer to *A-san no kao* 'A's face', *B-san no kao* 'B's face', and *C-san no kao* 'C's face' collectively as *mittu no kao* 'three faces'. This indicates that the "someone" referred to in "someone's face" is not a parameter. Regarding (41b), if a jar lid, a pan lid, and a frying pan lid were all placed on a table, one could say, "There are three lids on the table". With (41c) as well, someone who had to wash two windows in room A, one window in room B, and three windows in room C could say, "Today I have to wash six windows". According to this test as well, then, body-part nouns like *kao* 'face' and whole-part relation nouns like *huta* 'lid' and *mado* 'window' (archetypical inalienable possession nouns) are saturated nouns.

4.3 Situating unsaturated nouns in the lexical system

Several cautions regarding unsaturated nouns are called for. The first is that the distinction between saturated and unsaturated nouns cannot be handled simply as a lexical feature [±saturated]. In the case of unsaturated nouns in particular, important semantic constraints are imposed on their parameters. The semantic constraints are determined by each individual word so that, for instance, for the noun *syuyaku* 'leading role', both *eiga* 'movie' and *sibai* 'play' are possible values, but *sosuu* 'prime number', *mizu* 'water', *ringo* 'apple', or *yuki* 'snow' are not. Similarly, for *syain* 'member of a company', *kaisya* 'company' fills in the parameter but *e* 'picture' and *sibai* 'play' do not. Accordingly, the meaning differences among semantically similar unsaturated nouns like *kakite* 'writer, calligrapher', *sippitusya* 'author', *sakusya* 'author, creator', and *tyosya* 'author of a work', shown in (43), should probably be attributed to differences in their parameters.

(43) a. *kono {syoosetu / tyookoku / *ronbun / *kookyookyoku / *memo} no*
          this {novel   / sculpture / treatise / symphony   / memo} GEN
          sakusya
          author/creator

      b. *kono {*syoosetu / *tyookoku / ronbun / *kookyookyoku / memo} no*
          this {novel   / sculpture / treatise / symphony   / memo} GEN
          sippitusya
          author

      c. *kono {syoosetu / *tyookoku / ronbun / *kookyookyoku / *memo} no*
          this { novel   / sculpture / treatise / symphony   / memo} GEN
          tyosya
          author

Secondly, there are nouns that are ambiguous between being saturated nouns and unsaturated nouns. Generally, family relation nouns like *titi* 'father', *imooto* 'younger sister', *tuma* 'wife', *ko* 'child', *oba* 'aunt', and *itoko* 'cousin' are basically unsaturated nouns. Some family relation nouns, however, can be ambiguous. For instance, *kyoodai* 'siblings' in (44a), which refers to the concept 'sibling', is an unsaturated noun. In (44b), on the other hand, *ano kyoodai* 'those siblings' alone indicates a particular referent, so it is considered a saturated noun.

(44) a. *Hanako ni wa kyoodai ga inai.*
Hanako DAT TOP siblings NOM non.exist
'Hanako has no siblings.'

b. *Ano kyoodai wa naka ga ii ne.*
those siblings TOP relation NOM good SFP
'Those siblings get along well, don't they.'

The third point to note is that the parameters of unsaturated nouns are not always noun phrases but, as shown in (45), clauses may also appear (Yamaizumi 2010; Nishikawa 2013).

(45) a. [*Hanako ga rikon-si-ta*] *riyuu*
Hanako NOM get.divorced-do-PST reason
'the reason that Hanako got divorced'

b. [*Tyuuoo-sen ga okure-ta*] *gen'in*
Chuo-line NOM be.delayed-PST cause
'the cause of the Chuo-line being delayed'

c. [*tabako o kat-ta*] *oturi*
cigarettes ACC buy-PST change
'the change from buying cigarettes'

d. [*Taroo ga manbiki-si-ta*] *syooko*
Taro NOM shoplift-do-PST proof
'proof that Taro shoplifted'

e. [*hito o damasi-ta*] *batu*
people ACC deceive-PST punishment
'punishment for having deceived people'

Interestingly, the head nouns in (45), *riyuu* 'reason', *gen'in* 'cause', *oturi* 'change', *syooko* 'proof', and *batu* 'punishment' are all what Teramura (1992: 167) and Okutsu (1974) characterized as "nouns requiring a relative reference point" (termed "nouns of relativity"). The portions of the examples in (45) enclosed in brackets are called

sōtaiteki hojūsetsu 'relative supplementary clauses', and the relation between a relative supplementary clause and its head noun is undeniably an external relation. As shown in (46), the same relation holds between nouns showing temporal and spatial relations and the clauses that modify them.

(46) a. [*siai   ga   hazimar-u*] *mae*
        match NOM start-PRS   before
        'before the match starts'

   b. [*senkyo   ga   owat-ta*] *ato*
        election NOM end-PST   after
        'after the election ends'

   c. [*hukazake   o   si-ta*] *yokuzitu*
        drink.heavily ACC do-PST next.day
        'the day after a night of heavy drinking'

   d. [*daitooryoo ga   suwatte iru*] *migi*
        president   NOM sitting   is   right
        'to the right of where the president is sitting'

   e. [*Humiko   ga   tatte   iru*] *tonari*
        Fumiko NOM standing is   next.to
        'next to where Fumiko is standing'

From this, it looks as though we can consider what Teramura termed "nouns of relativity" to be unsaturated nouns taking clauses as their parameter values (Yamaizumi 2010). This opens up the interesting possibility of recasting what has been regarded as "external relation" noun modification as unsaturated nouns taking clauses as their parameter values.

4.4 Syntactic constructions in which unsaturated nouns play a role

The most important characteristic differentiating unsaturated nouns from other similar concepts is that the concept of unsaturated nouns plays a definitive role in determining the possibility of a certain type of construction. This section will examine this point. First, consider the following examples.

(47) a. *Ano hito   wa   syatyoo   da.*   (predicational sentence)
        that person TOP president COP
        'He is the president of the company.'

b. *Ano hito ga syatyoo da.* (specificational sentence)
  that person NOM president COP
  'It is he who is the president of the company.'

c. *Syatyoo wa ano hito da.* (inverted specificational sentence)
  president TOP that person COP
  'The president of the company is that person.'

All of the sentences in (47) are copular sentences composed of a subject and a predicate linked by a copula, but their meanings are quite different. Sentence (47a), where *ano hito* 'that person' is marked with the topic marker *wa*, is a sentence attributing the attribute *syatyoo* 'company president' to the referent of *ano hito* 'that person' and is called a "predicational sentence" in Nishiyama (2003). In this case *syatyoo* 'company president' functions as a predicate nominal. (47b), in which *ano hito* is marked with *ga*, shows the meaning "if you ask who the company president is, it is that person" and is called a specificational sentence in Nishiyama (2003). In (47b) *syatyoo* 'company president' shows the propositional function *X ga syatyoo da* 'X is the company president' and functions as "an NP involving a variable" (henceforth "NPIV") with the referent of *ano hito* 'that person' serving as the NP providing a value to fill in for the variable X (Kanbayashi 1988; Nishiyama 2003, 2008). Example (47c) is an inverted form of (47b) and has substantially the same meaning.

Since English does not, on the surface, have the distinction shown by *wa* and *ga* in Japanese, the English example (48a) can be read both as corresponding to (47a) and as corresponding to (47b) and is, in that sense, ambiguous (cf. Akmajian 1979; Higgins 1979; Declerck 1988; Nishiyama 2008).

(48) a. That person is the president.

b. The president is that person.

In the analysis of Higgins (1979), "specificational sentences" refer to the type of sentence in which NPIV (e.g. *the president*) rather than its value (e.g. *that person*) appears before the *be*-verb, as in (48b), but Moro (1997), among others, argues that this should be considered a copular sentence that has undergone inversion. If that is the case, then the Japanese sentence (47b) can be seen as a non-inverted type copular (specificational) sentence and (47c) as being derived from the basic form in (47b) through inversion.

Example (49) is of the pattern "*B (Hiroshima) ga A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*", but is actually a specificational sentence meaning "if you ask where the place of origin of oyster cuisine is, it is Hiroshima".

(49) *Hirosima ga kaki-ryoori no honba da.*
  Hiroshima NOM oyster-cuisine GEN place.of.origin COP
  'Hiroshima is the home of oyster cuisine.'

Here *kaki-ryoori no honba* shows the propositional function "*X* is the place for oysters" and *Hirosima* is the NP that provides a value for the variable *X*. The semantic structure of (49) is more clearly illustrated in (50).

(50) *Hirosima   ga   [kaki-ryoori   no   honba]   da.*
      Hiroshima NOM oyster-cuisine GEN place.of.origin COP

    Value NP   **NPIV**

    Hirosima   **[**X *ga*   *kaki-ryoori no honba da***]**

Note here that *honba* 'place of origin' is an unsaturated noun and *kaki-ryoori* 'oyster cuisine' is its parameter. (49) is of the form "*B (Hiroshima) ga A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*", but corresponding to (49) is sentence (51) of the form "*A (kaki-ryoori) wa, B (Hiroshima) ga C (honba) da*".

(51) *Kaki-ryoori   wa   Hirosima   ga   honba   da.*
      oyster-cuisine TOP Hiroshima NOM place.or.origin COP
      'For oyster cuisine, Hiroshima is the place.'

   Sentence (51) is closely related to (49) both syntactically and semantically. That is, (51) retains the basic meaning of (49). Specifically, (i) the fact that *honba* 'place of origin' and *kaki-ryoori* 'oyster cuisine' are related as unsaturated noun and associated parameter, (ii) the fact that *honba* is (part of) an NP involving a variable, (iii) the fact that Hiroshima shows the value that fills in that variable, and (iv) the fact that the sentence has the semantic structure of a specificational sentence in *Hirosima ga (sono) honba da* 'it is Hiroshima that is that place of origin' are all exactly the same as in (49). However, it is worth noting that, unlike (49), the overall structure of (51) is that of a predicational sentence with *kaki-ryoori* 'oyster cuisine' set up as topic and *Hirosima ga (sono) honba da* 'Hiroshima is its place of origin' showing an attribute that is attributed to it. In the end, the overall structure of (51) is that of a predicational sentence, but the predicate portion expressing the attributes (*Hirosima ga (sono) honba da* 'Hiroshima is its place of origin') itself has the structure of a specificational sentence, resulting in a "specificational clause – embedding predicational sentence" (cf. Nishikawa 2013: 199). Sentences like (51) are called "*kaki-ryoori* [oyster cuisine] sentences". Put more generally, let us call a sentence of the form "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*" a specificational sentence with the interpretation "the value of *X* that satisfies [*X ga A (kaki-ryoori) no C (honba) da*] is *B (Hirosima)*". When there is a specificational clause – embedding predicational sentence of the form "*A (kaki-ryoori) wa, B (Hiroshima) ga C (honba) da*" corresponding to this, we call the latter sentence a *kaki-ryoori* [oyster cuisine] sentence.

It should be noted that not all specificational sentences of the form "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*" have corresponding "*A (kaki-ryoori* 'oyster cuisine'*) wa, B (Hiroshima) ga C (honba* 'place of origin'*) da*" *kaki-ryoori* [oyster cuisine] sentences. For example, we do not have (52b) and (53b) as *kaki-ryoori* sentences corresponding to (52a) and (53a) (cf. Nishiyama 2003: 262).

(52) a. *Tama   ga   Isono-ke   no   neko   da.*   (Yamaizumi 2010: 14)
    Tama   NOM Isono-family GEN   cat   COP
    'It is Tama that is the Isono family's cat.'

  b. ?*Isono-ke   wa   Tama   ga   neko   da.*   (Yamaizumi 2010: 14)
    Isono-family TOP   Tama   NOM   cat   COP
    'As for the Isono family, it is Tama that is the cat.'

(53) a. *Taroo   ga   Bungakuza   no   haiyuu   da*
    Taro   NOM   Bungakuza.theater GEN   actor   COP
    'It is Taro that is an actor at the Bungakuza Theater.'

  b. ?*Bungakuza   wa   Taroo   ga   haiyuu   da.*
    Bungakuza.theater TOP   Taro   NOM   actor   COP
    'As for the Bungakuza Theater, it is Taro that is an actor.'

The question of under what circumstances a *kaki-ryoori* sentence "*A (kaki-ryoori* 'oyster cuisine'*) wa, B (Hiroshima) ga C (honba* 'place of origin'*) da*" corresponding to a specificational sentence "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*" is possible has been vigorously debated (Noda 1981, 1996; Nishiyama 1990, 2003; Kikuchi 1997). Concerning this question, Nishiyama (2003: 276) claimed that it is a necessary condition for the realization of a corresponding *kaki-ryoori* sentence "*A (kaki-ryoori* 'oyster cuisine'*) wa, B (Hiroshima) ga C (honba* 'place of origin'*) da*" corresponding to the specificational sentence "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*" that *C* be an unsaturated noun and that *A* show the value of its parameter. Since, in fact, there is the relation of an unsaturated noun and the value of its parameter between the *honba* of *C* and the *kaki-ryoori* of *A*, (51) is an acceptable *kaki-ryoori* sentence.

The reason that (52b) is not an acceptable *kaki-ryoori* sentence is that the word corresponding to C *neko* 'cat' is a saturated noun and, accordingly, there is no relation of an unsaturated noun and its parameter value between C *neko* 'cat' and A *Isono-ke* 'Isono family'. In the same way, since *haiyuu* 'actor' corresponding to C in (53a) is a saturated noun, the genitive phrase *Bungakuza no* and the head noun *haiyuu* are not in the relation "parameter + *no* + unsaturated noun" and hence, sentence (53b) is unacceptable as a *kaki-ryoori* sentence corresponding to (53a).

On the other hand, since there is a semantic relation of unsaturated noun and its parameter between *kokyoo* 'hometown' corresponding to C in (54b) and the *Taroo* filling the A slot, and between the *sidookyoozyu* 'academic advisor' and *Hanako* in the A position in (55b), as well as between the *syuyaku* 'leading role' in the C slot or (56b) and the *kono sibai* 'this play' in the A slot, they are all acceptable as *kaki-ryoori* sentences.

(54) a. *Aomori ga   Taroo no   kokyoo   da.* (specificational)
        Aomori NOM Taro   GEN hometown COP
        'It is Aomori that is Taro's hometown.'

   b. *Taroo wa   Aomori ga   kokyoo   da.* (*kaki-ryoori*)
        Taro   TOP Aomori NOM hometown COP
        'As for Taro, it is Aomori that is his hometown.'

(55) a. *Yamada-kyoozyu   wa   Hanako no   sidookyoozyu   da.*
        Yamada-professor TOP Hanako GEN academic.advisor COP
        'It is Prof. Yamada that is Hanako's academic advisor' (specificational)

   b. *Hanako wa   Yamada-kyoozyu   ga   sidookyoozyu   da.*
        Hanako TOP Yamada-professor NOM academic-advisor COP
        'As for Hanako, it is Prof. Yamada that is her academic advisor.' (*kaki-ryoori*)

(56) a. *Taroo ga   kono sibai no   syuyaku   da.* (specificational)
        Taro   NOM this   play GEN leading.role COP
        'It is Taro that has the lead in this play.'

   b. *Kono sibai wa   Taroo ga   syuyaku   da.* (*kaki-ryoori*)
        this   play TOP Taro   NOM leading.role COP
        'As for this play, it is Taro that has the lead role.'

How about a sentence like (57), then, in which the relation between the *kaminoke* 'hair' corresponding to C and the *Taroo* corresponding to A is one of "body-part noun and its base expression"? Note that, although there is nothing wrong with the specificational sentence (57a), the acceptability of the corresponding *kaki-ryoori* sentence (57b) is degraded.

(57) a. *Kore ga   Taroo no   kaminoke da.*   (B *ga* A *no* C *da*)
        this   NOM Taro   GEN hair   COP
        'It's this that is Taro's hair.'

   b. ?*Taroo wa   kore ga   kaminoke da.*   (A *wa*, B *ga* C *da*)
        Taro   TOP this   GEN hair   COP
        'As for Taro, it is this that is his hair.'

These observations demonstrate (i) that the concept of unsaturated nouns is a concept essential to stating the conditions for making *kaki-ryoori* sentences, and (ii) that body-part nouns like *kaminoke* 'hair' do not require parameters and are not unsaturated nouns.

Above, it was noted that Nishiyama (1990, 2003) stated that it was a necessary condition for the realization of a *kaki-ryoori* sentence "*A (kaki-ryoori* 'oyster cuisine'*) wa, B (Hiroshima) ga C (honba* 'place of origin'*) da*" corresponding to the specificational sentence "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*" that C be an unsaturated noun, and that A show the value of its parameter, but it should be emphasized that while this is a necessary condition, it is not a sufficient condition. That is, it is not the case that an acceptable *kaki-ryoori* sentence can always be formed when there is an unsaturated noun and its parameter relation between A and C. In fact, there are examples for which, even though in a *kaki-ryoori* sentence "*A (kaki-ryoori* 'oyster cuisine'*) wa, B (Hiroshima) ga C (honba* 'place of origin'*) da*" corresponding to the specificational sentence "*B (Hiroshima) ga, A (kaki-ryoori* 'oyster cuisine'*) no C (honba* 'place of origin'*) da*", C and A may be in an "unsaturated noun and its parameter" relation, the acceptability of the *kaki-ryoori* sentence is low. The examples in (58), (59), and (60) are from Sugioka and Kageyama (2011: 228).

(58) a. *Ziroo   ga   Yosida-sensei   no   musuko   da.*
        Jiro   NOM   Yoshida-teacher   GEN   son   COP
        'It is Jiro that is Prof. Yoshida's son.'

   b. **Yosida-sensei   wa,   Ziroo   ga   musuko   da.*
        Yoshida-teacher   TOP   Jiro   NOM   son   COP
        'As for Prof. Yoshida, it is Jiro that is his son.'

(59) a. *Hanako   ga   Ziroo   no   tomodati   da.*
        Hanako   NOM   Jiro   GEN   friend   COP
        'It is Hanako that is Jiro's friend.'

   b. ??*Ziroo   wa,   Hanako   ga   tomodati   da.*
        Jiro   TOP   Hanako   NOM   friend   COP
        'As for Jiro, it is Hanako that is his friend.'

(60) a. *Kore   ga   seikenkootai   no   kekka   da.*
        this   NOM   change.of.government   GEN   result   COP
        'It is this that is the result of a change of administration.'

   b. ??*Seikenkootai   wa,   kore   ga   kekka   da.*
        change.of.government   TOP   this   NOM   result   COP
        'As for the change in administration, it is this that is the result.'

Certainly the acceptability of (58b), (59b), and (60b) is low, but whether this low acceptability is due to semantic reasons or whether it is because there are pragmatic factors at work needs serious consideration. Whatever the necessary conditions for the formation of *kaki-ryoori* sentences may be, sufficient conditions are still not well established.

The above observations demonstrate that, among the complement-taking nouns that have been covered by the terms "inalienable possession nouns", "relational nouns", "relative nouns", and one-place nouns", the concept of "unsaturated nouns" stands out in particular and is deeply involved in the formation of particular constructions in Japanese.

5 Conclusion and future research perspectives

This chapter has taken up "complement-taking nouns" as a special group of nouns that exhibit different behavior from canonical nouns that denote entities and concepts by themselves. Their distinct property comes from their defective lexical semantic structure, which is not complete in isolation and requires for a complete interpretation the formation of a relation with other elements. There are many variations within this type of noun, but, taking the noun in question as NP2 and the other element (noun) that is essential to its interpretation as NP1, some things they all have in common include: (i) in the genitive construction "NP1 *no* NP2", the connection between NP1 and NP2 is not happenstance but is a necessary one; (ii) the meaning of NP2 is set in relation to NP1; (iii) the referent of NP2 only begins to function when combined with the referent of NP1; and (iv) the existence of the referent of NP2 presupposes the existence of the referent of NP1. Because of these factors, in terms of determining the meaning of NP2, NP1 is a required argument and is referred to variously as an "argument required by NP2", a "variable", a "complement", or a "parameter". The categories of nouns treated under the rubrics of "inalienable possession nouns", "relational nouns", "relative nouns", "one-place nouns", "unsaturated nouns", and "nouns with relativity" in studies of Japanese are precisely this type of noun.

In this chapter, we have examined similarities and dissimilarities among these categories to disentangle the complex relations among them. Notwithstanding the fact that there is a large overlap in the nouns given as examples of each of these categories, the approaches from which each category concept defines the semantic relation of NP1 *no* NP2 differ subtly. In particular, the distinction between the archetypical inalienable possession nouns like body-part nouns (*kaminoke* 'hair', *asi* 'feet', *kata* 'shoulder, and the like) and whole-part relation nouns (*huta* 'lid', *genkan* 'front entrance', *yane* 'roof', *hyoosi* 'cover sheet', and the like), on the one hand, and unsaturated nouns (*zyoosi* 'boss', *imooto* 'younger sister', *honba* 'place of origin',

riyuu 'reason', and the like), on the other, is very important. Unsaturated nouns are a type of noun that include in their meaning a parameter "*X no* [of X]" and, unless a value for the parameter is set, the extension of the noun cannot be determined, but body-part nouns and whole-part relation nouns do not share this characteristic. The distinction between the concept of unsaturated nouns and notions of body-part and whole-part relation nouns is also important with regard to constraints on the formation of certain constructions. With regard to the *kaki-ryoori* construction, unsaturated nouns and body-part nouns show remarkable differences. Also, the notion of unsaturated nouns opens up the possibility of a new analysis for what Teramura (1992) called "external relation" noun modification structures like [*Hanako ga rikonsita*] *riyuu* 'the reason that Hanako got divorced' by extending the domain of parameter values from noun phrases to clauses. On the other hand, the notion of "one-place nouns" as a text level phenomenon is extremely important for explicating cohesiveness in particular.

It should be stressed that the notions of "complement-taking nouns" in general, and that of "unsaturated nouns" in particular, are not just a matter of lexical semantics but are deeply involved in the system of Japanese grammar as an interface between the lexicon and syntax. Among the unsolved issues, only two will be singled out here. One is to deepen the understanding of their nature from the viewpoint of the whole grammatical system. The exposition in this chapter has been wholly descriptive and taxonomic. It is necessary to consider how the syntactic behavior of unsaturated nouns can be formally captured in their lexical representations and how it is linked to syntactic structure. The other topic for future research is the assessment of the proposed notions in the world's languages. At present, we do not have sufficient data to test whether the same arguments for distinguishing several types of "inalienability" and pinpointing the class of unsaturated nouns are available in other languages. No matter whether the class of unsaturated nouns turns out to be universally available or specific to Japanese, we must ask exactly where their intriguing properties come from.

Acknowledgments

References

Akmajian, Adrian. 1979. *Aspects of the grammar of focus in English*. New York: Garland.
Barker, Chris. 2011. Possessives and relational nouns. In Claudia Maienborn, Klaus von Heusinger and Paul Portner (eds), *Semantics: An international handbook of natural language meaning*, 1109–1130. Berlin: Mouton de Gruyter.
Chomsky, Noam. 1972. *Studies on semantics in generative grammar*. The Hague: Mouton.

Croft, William. 1990. *Typology and universal*. Cambridge: Cambridge University Press.

Cruse, D. Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Declerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo clefts*. Leuven: Leuven University Press.

Dixon, Robert Malcolm Ward. 2009. *Basic linguistic theory: Volume 2: Grammatical topics*. Oxford: Oxford University Press.

Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.

Higgins, Francis Roger. 1979. *The pseudo cleft construction in English*. New York: Garland.

Iori, Isao. 2007. *Nihongo ni okeru tekisuto no kessokusei no kenkyū* [A study of cohesion in Japanese text]. Tokyo: Kurosio Publishers.

Kanbayashi, Yoji. 1988. Soteibun to shiteibun: *wa* to *ga* no ichimen [Predicational sentences and specification sentences: Some properties of *Wa* and *ga*]. *Tsukuba Daigaku Bungei Gengo Kenkyū Gengo hen* [Annual of Tsukuba University: Studies in Language and Literature (Language)] 14. 57–74.

Kikuchi, Yasuto. 1997. 'Kakiryori wa Hiroshima ga honba da' kōbun no seiritsu jōken [Well-formedness conditions of 'kakiryori wa Hiroshima ga honba da' construction]. *Hiroshima Daigaku Nihongo kyōiku gakka Kiyō* [Bulletin of the Department of Teaching Japanese as a Second Language, Hiroshima University] (A special issue celebrating the 10th anniversary of the founding of the department) 7. 89–107.

Moro, Andrea. 1997. *The raising of predicates: Predicate noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.

Nichols, Johanna. 1988. On alienable and inalienable possession. In William Shipley (ed.), *In honor of Mary Haas: From the Haas festival conference on native American linguistics*, 557–609. Berlin: Mouton de Gruyter.

Nishikawa, Kenya. 2013. 'NP$_1$ no NP$_2$' taipu F: Jōto-fukanō-meishi NP$_2$ to sono kitai hyōgen NP$_1$ ['NP$_1$ no NP$_2$' Type F: Inalienable noun NP$_2$ and its base expression NP$_1$]. In Yuji Nishiyama (ed.), *Meishiku no sekai: Sono imi to kaishaku no shinpi ni semaru* [On noun phrases: The mystery of meaning and interpretation], 65–82. Tokyo: Hituzi Syobo.

Nishiyama, Yuji. 1990. 'Kakiryōri wa Hiroshima ga honba da' kōbun ni tsuite: Hōwa-meishiku to hihōwa-meishiku [On 'kakiryōri wa Hiroshima ga honba da' construction: Saturated NPs and unsaturated NPs]. *Keiōgijyuku daigaku Gengo bunka Kenkyūsho Kiyō* [Reports of the Keio Institute of Cultural and Linguistic Studies] 22. 169–188.

Nishiyama, Yuji. 2003. *Nihongo meishiku no imiron to goyōron: Shijiteki meishiku to hi shijiteki meishiku* [Semantics and pragmatics of noun phrases in Japanese: Referential noun phrases and non referential noun phrases]. Tokyo: Hituzi Syobo.

Nishiyama, Yuji. 2008. Non-referentiality in certain noun phrases. In Tetsuya Sano, Mika Endo, Miwa Isobe, Koichi Otaki, Koji Sugisaki, and Takeru Suzuki (eds.), *An enterprise in the cognitive science of language: A festschrift for Yukio Otsu*, 13–25. Tokyo: Hituzi Syobo.

Noda, Hisashi. 1981. 'Kakiryori wa Hiroshima ga honba da' kōbun ni tsuite [On *kakiryori wa Hiroshima ga honba da* construction]. *Machikaneyama Ronsō Nihongakuhen* (School of Letters, Osaka University) 16. 45–66.

Noda, Hisashi. 1996. '*Wa*' to '*ga*' [*Wa* and *ga*]. *Sin nihongo bunpō sōsho* 1 [New series of Japanese grammar 1]. Tokyo: Kurosio Publishers.

Okutsu, Keiichiro. 1974. *Seisei nihongo bunpō ron: Meishiku no kōzō* [Generative Japanese grammar: The structure of noun phrases]. Tokyo: Taishukan Shoten.

Ono, Naoyuki. 2009. Nihongo rentai-shūshoku-setsu e no goi-imiron-teki apurōchi [A lexical-semantic approach to Japanese adnominal clauses]. In Yoko Yumoto and Hideki Kishimoto (eds.), *Goi no imi to bunpō* [Lexical meaning and grammar], 253–272. Tokyo: Kurosio Publishers.

Sugioka, Yoko and Taro Kageyama. 2011. Meishi-ka to kō no uketsugi [Argument inheritance in nominalization]. In Kageyama Taro (ed.), *Nichi ei taishō: meishi no imi to kōbun* [The semantics and syntax of nouns: A contrastive study of English and Japanese], 209–239. Tokyo: Taishukan Shoten.

Teramura, Hideo. 1992 [1975]. Rentai-shūshoku no shintakusu to imi [The syntax and semantics of noun modification] 1–4. In *Teramura Hideo ronbun shū I: Nihongo bunpō hen* [Collected papers of Hideo Teramura I: Japanese grammar], 157–320. Tokyo: Kurosio Publishers.

Tsunoda, Tasaku. 1995. Possession cline in Japanese and other Languages. In Hilary Chappell (ed.), *The grammar of Inalienability*, 565–630. Berlin: Mouton de Gruyter.

Tsunoda, Tasaku. 2009. *Sekai no gengo to nihongo (Kaitei ban): Gengo ruikei ron kara mita nihongo* [The world's languages and Japanese (Revised edition): A typological perspective on Japanese language.] Tokyo: Kurosio Publishers.

Yamaizumi, Minoru. 2010. *Setsu ni yoru hihōwa meishi(ku) no paramēta no hojū* [Unsaturated nouns in Japanese: How clauses fill their parameters]. Tokyo: University of Tokyo dissertation.

Yamaizumi, Minoru. 2013. Hihōwa meishi to sono paramēta no atai [Unsaturated nouns and values of their parameters]. In Yuji Nishiyama (ed.), *Meishiku no sekai: Sono imi to kaishaku no shinpi ni semaru* [On noun phrases: The mystery of meaning and interpretation], 11–27. Tokyo: Hituzi Syobo.

Hideki Kishimoto
19 Idioms

1 Introduction

Idioms are syntactic phrases, like *pull someone's leg* and *let the cat out of the bag*, that have unique properties that crosscut the lexicon and syntax. From a semantic point of view, their meanings are not constructed by putting together the meanings of their parts compositionally. Strictly speaking, the degree of semantic opacity differs in individual idioms, ranging from fully opaque idioms like *pull someone's leg* 'tease someone' and *kick the bucket* 'die' to moderately opaque idioms like *let the cat out of the bag* 'reveal a secret' and *lend an ear* 'listen carefully', where the noun components hint at certain metaphorical or metonymic meanings stemming from them, such as 'a secret' for the noun *cat* in *let the cat out of the bag* and 'auditory attention' for the noun *ear* in *lend an ear*.[1] Japanese has a wide variety of fully and moderately opaque idioms. From a syntactic vantage point, idioms are composed of syntactically transparent internal constituents, and yet re-arrangement of their internal composition by syntactic operations such as passivization is, in many cases, impossible or results in the loss of idiomatic meaning. From a morphological viewpoint, the component words that make up a particular idiom are limited to lexically designated items. The noun *leg* in *pull someone's leg*, for example, cannot be replaced with *foot* or *arm* on the relevant meaning as an idiom. Because of these idiosyncratic properties not shared with ordinary syntactic phrases, idioms need to be registered in the lexicon as a whole.

While Japanese abounds with idioms consisting of predicates and a variety of nouns functioning as their arguments, there are not many works in the literature that directly address the theoretical issues surrounding them. This chapter attempts to delineate notable properties of Japanese idioms, with particular attention to those involving body-part nouns (or nouns related to human attributes) such as *mi ni tukeru* 'master, acquire (a skill)' and *mimi o kasu* 'lend an ear, listen carefully'. Since Japanese exploits case particles to represent arguments, it is appropriate to distinguish several different types of idioms according to the transitivity and case frames of the predicates involved, as shown in Table 1.

---

1 It is worth noting that idioms tend to have fixed inflectional forms. For instance, in an idiom like *pull someone's leg*, the body part noun leg must appear in the singular form, but not in the plural form (**pull someone's legs*).

Table 1: A classification of idioms according to case frames

Transitive idioms [NP$_{ACC/DAT}$ V]	*koe o kakeru* 'invite', *te ni ireru* 'get'
Ditransitive idioms [NP$_{DAT}$ NP$_{ACC}$ V]	*kyuusi ni issyoo o eru* 'narrow escape'
Unaccusative idioms [NP$_{NOM}$ V$_{UNACC}$]	*gata ga kuru* 'come close to a breakdown'
Passive idioms [NP V-Pass]	*akke ni tor are ru* 'be taken aback'
Negative idioms [NP V-Neg]	*rati ga aka nai* 'get nowhere'
Possessor-raising Idioms [X$_{\{GEN/DAT\}}$ NP V]	X *{no/ni} kao ga kiku* 'have a pull'

The discussion in this chapter proceeds as follows. Section 2 surveys the general properties of idioms and Section 3 discusses varying degrees of morphosyntactic fixedness. In the following three sections, some characteristic properties of Japanese idioms are discussed, with particular attention to three types of idioms: unaccusative and passive idioms (Section 4), possessor-raising idioms (Section 5), and negative idioms (Section 6). The conclusion is presented in Section 7.

2 General features of Japanese idioms

Idioms convey idiosyncratic meanings that cannot be retrieved directly from the meanings of their parts.[2] As an example, let us consider the idiom *me o mukeru* 'get

---

2 There are multifarious views on what expressions should count as idioms, and the precise defini tion of idioms differs significantly from one proposal to another (see e.g. Everaert et al. (eds.) 1995; Nunberg, Sag and Wasow 1994). This chapter follows one of the most restricted views, taking idioms as syntactically complex predicative expressions with non compositional meanings. At first sight, idioms might look similar to fixed expressions such as proverbs, quotes, and maxims (e.g. *Inu mo aruke ba boo ni ataru* 'Everyone has good days', *Isi no ue ni mo san nen* 'Patience wins in the end'), as well as titles, names, foreign phrases. In principle, the fixed expressions mentioned here are usable only as a whole, and neither modification nor embedding (as part of a clausal constituent) is allowed, as in (i). (# indicates that the sentence is not interpretable in the idiomatic meaning.)

(i) a. *#Inu mo   aruke ba boo   ni   atari hazime ru.*
        dog   also walk if   stick LOC hit begin PRS
        'If a dog walks, she begins to hit a stick (≠ Everyone has good days).'

     b. *#Isi   no   ue   ni   mo   san nen   suwari soo   da.*
        stone GEN top LOC also three year sit   likely COP
        'It is likely that I will sit on a stone for three years (≠ It is likely that someday patience will win in the end).'

In contrast, the syntactic behavior of idioms is governed by regular rules, and they can be modified or allow embedding, accordingly. Subject idioms (clausal idioms), such as *kumo yuki ga ayasiku naru* 'Matters take a turn for the worse' and *kankodori ga naku* 'there are less and less customers', at first sight look similar to proverbs, since the clauses carry idiomatic meanings, but differ from them in allowing syntactic operations such as embedding, as in (ii).

interested in' (as found in an expression like *kaigai ni me o mukeru* [abroad LOC eye ACC turn] 'get interested in the world abroad'). When interpreted literally, the VP sequence of *me o mukeru* has a lexical meaning like 'X causes Y's eye to become turned to Z', but the idiom is actually related to a non-literal lexical meaning like 'X gets interested in Y', as indicated in (1).

(1)   [$_{VP}$ [$_{NP}$ *me o*] *mukeru*]

 'X gets interested in Y'   ←   'X causes Y's eye to become turned to Z'

The idiomatic meaning of 'having an interest' does not come from the meanings of their parts, because neither *me* 'eye' nor *mukeru* 'turn(tr.)' denotes 'interest'.[3] Even though *mukeru* is a transitive predicate taking an object, the idiom allows only the body-part noun *me* 'eye' to appear in the position where the object is supposed to appear, which suggests that in the idiom, the slot where a variable normally appears

---

(ii) a. *Kumo yuki   ga   ayasiku   nari hazime ta.*
          cloud   NOM doubtful   become begin PST
          'Matters began to take a turn for the worse.'

  b. *Ituka   kankodori ga   naki soo   da.*
          someday cuckoo   NOM sing likely   COP
          'It is likely that someday, there will be less and less customers.'

It should also be noted that Japanese idioms often allow insertion of adverbial particles within their chunks, even if some syntactic operations such as passivization are blocked. Jackendoff (1997) take clichés to have 'nothing non compositional in their syntax and meaning' (p. 155), and thus distin guishes them from idioms.

3 It is conceivable that idioms start out from figurative expressions. Nevertheless, this does not mean that figurative expressions automatically qualify as idioms. For instance, the expression *utagai no me o mukeru* [suspect GEN eye ACC turn(tr.)] 'suspect' imposes figurative (or metaphoric) interpre tation on the body part noun *me* 'eye', and the sequence of the body part noun and the predicate looks like it forms an idiom at first brush. But there is good reason to believe that this expression does not count as an idiom. For one thing, the body part expression *utagai no me* can be associated with a variety of predicates, as in *utagai no me de miru* [suspect GEN eye INST look.at] 'look suspi ciously' and *utagai no me o motu* [suspect GEN eye ACC have] 'have a suspicion'. For another, the sequence of *utagai no me o mukeru* can easily be broken up via scrambling.

(i) a. *Ken wa sore ni   utagai no   me o   muke ta.*
          Ken TOP that LOC suspect GEN eye ACC turn(tr.) PST
          'Ken suspected that.'

  b. *Ken wa [utagai no   me o]$_i$ sore ni   t$_i$ muke ta.*
          Ken TOP suspect GEN eye ACC that LOC   turn(tr.) PST
          'Ken suspected that.'

For a certain sequence of words to be categorized as an idiom, the constituents need to be adjacent (although there are cases where certain modifiers are allowed)   a syntactic requirement that is often referred to as the "adjacency" condition (see Section 4). Owing to the adjacency condition, idioms do not allow idiomatic arguments to be scrambled across non idiomatic arguments. But this restriction

is filled by a constant, i.e. the designated body-part noun. Nevertheless, the idiom *me o mukeru* behaves as a verb phrase syntactically, with a transparent syntactic structure consisting of the object *me* 'eye' plus the verb *mukeru* 'turn(tr.)', for its constituents can undergo certain syntactic operations (e.g. this idiom can be nominalized, as in *me no muke-kata* [eye GEN turn(tr.)-way] 'the way of getting an interest', where accusative *o* inside the idiom constituency is changed to genitive *no* by virtue of the fact that the predicate is nominalized with the addition of the suffix *-kata* 'way').

The overall behavioral patterns of Japanese idioms have been extensively discussed by Miyaji (ed.) (1982) and Miyaji (1999). Miyaji (ed.) (1982) reports that in his list of approximately 1200 idioms (or more than 1400 idioms if transitive/intransitive pairs are counted separately), verb-based idioms like (2a) account for 63% of the listed idioms, whereas adjective-based idioms like (2b) account for 15%. There are also a smaller number of noun-based idioms as well as other types of idioms, such as (2c) and (2d).

(2)　a.　*me o mawasu* [eye ACC turn(tr.)] 'pass out'

　　b.　*hara ga kuroi* [stomach NOM black] 'malicious'

　　c.　*ato no maturi* [post GEN festival] 'no use afterwards'

　　d.　*nemimi ni mizu* [sleeping.ear LOC water] 'sudden news'

Verb-based idioms show a variety of argument realization patterns, depending on the case frames required by individual verbs, as exemplified by (3).

(3)　a.　*keri o tukeru* [end ACC attach(tr.)] 'put an end to'

　　b.　*kuti ga tatu* [mouth NOM stand(intr.)] 'speak fluently'

　　c.　*te ni hairu* [hand LOC enter] 'be acquired'

　　d.　*ha ni kinu o kise-nai* [tooth DAT cloth ACC dress-NEG] 'not mince matters'

---

does not apply to the expression *utagai no me o mukeru*. Furthermore, a similar expression like *me o mukeru* 'get interested in' does not allow the body part noun to be scrambled over a non idiomatic argument, as in (ii).

(ii)　a.　*Ken　wa　kaigai　ni　me　o　muke ta.*
　　　　Ken　TOP　abroad　LOC　eye　ACC　turn(tr.) PST
　　　　'Ken got interested in the world abroad.'

　　b.　?*Ken　wa　[me　o]$_i$　kaigai　ni　t$_i$　muke ta.*
　　　　Ken　TOP　eye　ACC　abroad　LOC　　　turn(tr.) PST
　　　　'The world abroad, Ken got interested in.'

Here, the adjacency condition applies to *me o mukeru*. The crucial difference is that the verb *mukeru* 'turn(tr.)' is used in a literal sense in the case of *utagai no me o mukeru*, whereas the verb, as well as the body part noun, is used in a non literal sense in the case of *me o mukeru* 'get interested in'. In latter case, the entire sequence forms an idiom.

The example in (3a) represents an "accusative-V" idiom, where a transitive verb selects an accusative-marked object. Likewise, the examples in (3b) and (3c) respectively represent a "nominative-V" idiom and a "locative (or dative)-V" idiom. According to Miyaji (ed.) (1982), accusative-V idioms are the most common among verb-based idioms, accounting for 57%, whereas nominative-V idioms and locative/dative-V idioms represent 20% each. There are a small number of ditransitive idioms comprising both direct and indirect objects inside their idiom sequences, as in (3d).[4] In addition, no idioms are found in Japanese that consist of a transitive verb and its subject, excluding the object. Generally, the percentages are in accord with the claim about the tendency observed for English by Marantz (1984), who claims that the closer the argument is to the verb, the more easily it forms part of an idiom constituent.

The examples given in (3) all correspond syntactically to English idioms of the *shoot the breeze* type, where both the verb and the noun are fixed to lexically designated items. Japanese also has idioms of the *pull someone's leg* type, where the verb *pull* and the noun *leg* are lexically fixed but the possessor of *leg* is a variable that can be filled by any noun representing a human. Additional examples of such idioms are shown in (4).

(4) a. *X no senaka o osu* [X GEN back ACC push] 'give X a supportive push'

   b. *X no te ni hairu* [X GEN hand LOC enter] 'be acquired by X'

   c. *X no asi o hipparu* [X GEN leg ACC pull] 'stand in X's way'

In (4), *X* represents a variable embedded within the idiom sequence and this slot can be filled by any kind of regular expression, insofar as it satisfies the selectional restriction imposed by the host noun. In many cases, the host nouns are body-part nouns like *senaka* 'back' and *te* 'hand'. It is sometimes possible to realize *X* as a clausal argument by virtue of possessor raising (see Section 5).

Closely related to the non-compositionality of idiom interpretation noted above is the fact that idioms are constructed from a restricted or fixed choice of words. Since idioms are formed on lexically designated words, they generally do not allow their constituents to be replaced with synonymous words. For instance, when *abura* in *abura o uru* is replaced by another noun referring to oil, the idiom loses its idiomatic meaning, as shown in (5a), where the symbol "#" indicates that the expression is infelicitous on the idiom interpretation.

---

4 Verbs like *ireru* 'put in' and *tukeru* 'attach(tr.)' are three place predicates, which take a locative (or goal) argument, alongside a subject and an object. Since the locative argument behaves as a PP syntactically, this class of verbs is referred to as 'transitive' verbs, and idioms formed on this type of verbs are called 'transitive' idioms in this chapter (see Sections 3 and 5).

(5) a. *#{oiru/gasorin/tooyu} o uru* [{oil/gasoline/kerosene} ACC sell] 'idle away'

b. *X ni {piriodo/syuusihu} o utu* [X LOC {period/full.stop} ACC strike] 'put a period (i.e. and end) to X'

There are cases, however, where replacement with synonymous words is possible though to a limited extent. In (5b), both *piriodo* 'period' (Western loan word) and *syuusihu* 'period' (Sino-Japanese word) permit the relevant interpretation as an idiom. It is conceivable that the Sino-Japanese rendition might be due to a calque (or a loan translation) of the corresponding English idiom. In any event, as opposed to the verbs in idiom chunks, which stem from the native Japanese lexical stratum, the nouns that collocate with them may come not only from the native Japanese lexical stratum but also from the Sino-Japanese and foreign lexical strata, as illustrated in (6).

(6) a. *sazi o nageru* [spoon ACC throw] 'give up'   [native Japanese noun]

b. *kigen o toru* [mood ACC take] 'flatter'   [Sino-Japanese noun]

c. *sutoppu o kakeru* [stop ACC hang(tr.)] 'stop' [foreign noun]

One important feature of Japanese idioms, often noted in the literature (e.g. Miyaji 1999), is that inalienable body-part nouns are frequently combined with verbs to give idiom meanings. The body-part nouns most typically used in idioms include *ki* 'mind', *kokoro* 'heart', *te* 'hand', *me* 'eye', *kuti* 'mouth', *hana* 'nose', and *mi* 'body', as exemplified in (7).

(7) a. *te o ireru* [hand ACC put.in] 'revise'

b. *me ni hairu* [eye LOC enter] 'notice'

c. *mimi ni hairu* [ear LOC enter] 'come into one's attention'

d. *hana ni tuku* [nose LOC attach(intr.)] 'do not like'

e. *kokoro ni nokoru* [heart LOC remain] 'be impressed'

f. *kuti ga karui* [mouth NOM light] 'have a loose tongue'

g. *mi ni tukeru* [body LOC attach(tr.)] 'master, acquire (a skill)'

The component nouns appearing in idioms are non-referential, i.e. they do not refer to any particular entities, although their possessor may often be referential (see Section 5). More importantly, the body-part nouns used in idioms tend to represent metaphorical or metonymic meanings related to the mental or physiological functions of the body parts involved. *Mimi* 'ear' in (6c), for example, does not refer to a physical entity but to an auditory perception in acquiring information through the ears.

Some idioms include archaic or obsolete nouns that are not normally used today in other contexts. The nouns *udatu* and *hu* that appear in the idioms of (8a, b) are used to refer to 'roofed wall' and 'gut', respectively, but these meanings are no longer transparent in contemporary Japanese.[5]

(8) a. *udatu ga agara-nai* [roofed.wall NOM rise-NEG] 'not get a promotion'

    b. *hu ni oti-nai* [gut LOC drop-NEG] 'suspect'

Presumably, these obsolete words are remnants of expressions that were active at the time of idiom formation (cf. Brinton and Traugott 2005). On the other hand, some nouns have transparent meanings even if their occurrence is limited to idioms.

(9) a. *okubi ni mo dasa-nai* [neck LOC also put.out-NEG] 'not say a word'

    b. *okumen mo nai* [scared.face also NEG] 'not shamed at all'

    c. *kyuusi ni issyoo o eru* [9.death LOC one life ACC get] 'narrow escape'

Nouns such as *okubi* 'neck', *okumen* 'scared face', and *kyuusi* '9 deaths' are transparent in meaning, but are normally not used outside the domain of idioms.

There is also a tendency for the repertoire of verbs available for idioms to be lexically limited. The verbs commonly exploited in idioms include *tukeru/tuku* 'attach (tr.)/attach (intr.)', *kakeru* 'hang', *ireru* 'put in', *dasu* 'let out', and *toru* 'take'. These verbs are fairly productively associated with body-part nouns to derive idioms. The verb *tukeru* 'attach(tr.)', for example, can be combined with a variety of body-part nouns, as in (10).

(10) a. *me o tukeru* [eye ACC attach(tr.)] 'have an eye on'

    b. *kuti o tukeru* [mouth ACC attach(tr.)] 'have a bite, eat'

    c. *te o tukeru* [hand ACC attach(tr.)] 'start'

    d. *ki o tukeru* [mind ACC attach(tr.)] 'mind'

Combination of these verbs with body-part nouns, however, is not entirely free, with gaps occasionally found, as seen in the impossibility of the sequence *mimi o tukeru [ear ACC attach(tr.)] as an idiom.

There are cases where the possibility of idiom formation differs depending on the kind of case marking assigned to a noun, as in (11).

(11) *mi {ni/*o} tukeru* [body {LOC/*ACC} attach(tr.)] 'master, acquire (a skill)'

---

5 Although words like *hu* and *utdatu* are not normally used in ordinary contexts, they may still be used in some special contexts. In addition, some common fixed expressions are also found that contain these words, such as *hu nuke* [viscera fall.out] 'a coward'.

Unlike the idioms given in (10), where the body-part nouns are marked with accusative *o*, the idiom in (11) – formed by combining the body-part noun *mi* 'body' and the verb *tukeru* 'attach(tr.)' – is assigned the idiom interpretation when the noun is marked with locative *ni* rather than accusative *o*. The disparity in case marking between (10) and (11) can be explained from the syntactic and semantic properties of the verb *tukeru* 'attach(tr.)'. This verb calls for the case frame <NOM, LOC, ACC> in its original usage, and maintains this case frame even when used in the idiom chunks. For the idiom in (11), *mi* 'body' is marked with locative *ni* because it is conceived of as the place where some skill or knowledge is stored. In fact, the skill or knowledge acquired can be represented as an accusative object that is placed in front of the whole idiom, as in *gizyutu o mi ni tukeru* [skill ACC body LOC attach(tr.)] 'acquire a skill'. Conversely in the idioms of (10), the body-part nouns are marked with the accusative *o* because they are construed as the themes of attaching, so that the places to which they are metaphorically attached are represented with locative *ni*, as in *sono otoko ni me o tukeru* [that man LOC eye ACC attach(tr.)] 'have one's eye on that man' and *tabemono ni kuti o tukeru* [food LOC mouth ACC attach(tr.)] 'have a bite of the food'.

While idioms maintain the syntactic case frames, the idiom sequences often display behaviors that are radically different from those expected from their non-idiomatic counterparts. Observe the examples in (12).

(12)　a.　*kokyoo ni nisiki o kazaru* [home to brocade ACC decorate] 'return in glory'

　　　b.　*musizu ga hasiru* [insect.acid NOM run] 'get disgusted'

　　　c.　*mimi o kasu* [ear ACC lend] 'listen to'

The idiomatic and literal lexical meanings of the idioms in (12) are illustrated in (13).

(13)　a.　*kokyoo ni nisiki o kazaru*
　　　　　'X returns to X's hometown in glory' < 'X causes Y to become decorated on Z'

　　　b.　*musizu ga hasiru*
　　　　　'get disgusted' < 'gastric juice comes out'

　　　c.　Y *ni mimi o kasu*
　　　　　'X listens to Y' < 'X causes Y to possess Z (temporarily)'

Kazaru 'decorate' by itself is originally a causative verb of change of state, so the verb phrase created by combining the verb with an object denotes a change of state. In (13a), however, the VP sequence represents a motional event, but not a change of state, because it denotes a return to one's hometown after making a big success in life. The verb *hasiru* 'run' is an unergative verb denoting an intentional action and takes an agent as its subject when used in a non-idiomatic sense; nevertheless, the idiom in (13b) includes the inanimate subject *musizu* 'gastric juice'. Similarly, when

used non-idiomatically, the verb *kasu* 'lend' behaves as a double object verb select-
ing a human recipient for its *ni*-marked indirect argument, e.g. *{Ken/*sono hanasi} ni
hon o kasu* [{Ken/*that story} DAT book ACC lend] 'lend a book to {Ken/*that story}';
in (13c), on the other hand, the *ni*-marked argument represents an entity to which
attention is directed, e.g. verbal messages or the source which they come from, so
that this argument may be inanimate, e.g. *{Ken/sono hanasi} ni mimi o kasu* [{Ken/
that story} DAT ear ACC lend] 'listen to {Ken/that story}'.

   Some idioms select only verbs in the passive or causative forms. The examples
in (14) represent idioms with verbs in the causative *-(s)ase*.

(14) a. *haba o kik-ase-ru* [width ACC take.effect-CAUS PRS] 'has great influence'

   b. *kuti o togar-ase-ru* [mouth ACC stick.out-CAUS PRS] 'pout one's mouth'

   c. *hanagusuri o kag-ase-ru* [nose.medicine sniff-CAUS PRS] 'bribe'

The idiom interpretation disappears if the causative verbs are replaced by the corre-
sponding simple, non-causative verb forms. For example, the non-causative counter-
part of (14a), **haba ga kiku* [width NOM take.effect], makes no sense. In a similar
vein, the verbs used in the idioms in (15) must take the passive form.

(15) a. *akke ni tor-are-ru* [aback by take-PASS PRS] 'be taken aback'

   b. *ki o tor-are-ru* [mind ACC take-PASS PRS] 'be distracted'

   c. *usiroyubi o sas-are-ru* [finger ACC point-PASS PRS] 'be criticized behind
      one's back'

The non-passive counterpart of (15a) should be **akke {o/ni} toru* [aback {ACC/LOC}
take], which is deemed simply uninterpretable.

   Finally, it is instructive to see that a fairly large number of Japanese idioms
have corresponding compound words that carry the same idiom interpretations, as
exemplified in (16).

(16) a. *sewa*   *o*   *yaku*   >   *sewa-yaki*
      care   ACC   burn(tr.)   care-burn(tr.)
      'take care of'   'taking care of'

   b. *ki*   *o*   *tukau*   >   *ki-zukau*
      mind   ACC   use   mind-use
      'pay attention'   'pay attention'

   c. *yaku*   *ni*   *tatu*   >   *yaku-datu*
      use   LOC   stand(intr.)   use-stand(intr.)
      'be of use'   'be of use'

A plausible way to account for such correspondences in form and meaning would be to assume that the idiom structures are rendered into compound words via a lexical process of compounding. On this analysis, the compound noun *sewa-yaki* in (16a) is derived from the idiom sequence *sewa o yaku* via V-to-N conversion (i.e. *yaku* (V) > *yaki* (N)) and N+N compounding. In (16b, c), the two components of the noun + verb sequences are reduced to single words solely via compounding, i.e. *ki-zukau* and *yaku-datu*. The status of the reduced forms in (16) as compound words is confirmed by *rendaku* (or sequential voicing), a phonological process that applies to the second member of a compound word (e.g. *yaku ni **t**atu* > *yaku-**d**atu*). Note that the compound forms in (16b, c) function as verbal predicates, in the same way as their periphrastic sources. The productivity of the word formation rule giving rise to compound verbs in the form [$_V$ N+V] is not high, however. It is not the case that a compound verb can regularly be formed from a transitive verb phrase via compounding (e.g. **hon-yomu* [book-read] < *hon o yomu* [book ACC read]), but still, examples like *katati-zukuru* [shape-make] 'form' and *te-watasu* [hand-pass] 'hand' are found among idiomatic expressions (see Chapter 7 [Kageyama, this volume] for more details on N-V compounding).

3 Syntactic constituency of idioms

Idioms have different degrees of morphosyntactic fixedness, and thus show a significant difference in applicability of replacement or movement operations (see, e.g. Nunberg, Sag, and Wasow 1994; Fraser 1974; Culicover 1976; Everaert et al. (eds.) 1995). This facet of idioms has attracted much attention in previous studies on Japanese idioms. Some authors (e.g. Miyaji (ed.) 1982; Miyaji 1999; Asuka 1982; Ishida 2000, 2004) claim that there is a cline on the applicability of syntactic operations. In the literature, a number of different linguistic phenomena are discussed, some of which I will discuss below.

The component nouns in idioms show varying degrees of modifiability (Miyaji (ed.) 1982; Asuka 1982; Ishida 2000, 2004). For instance, some idioms like *te o utu* 'take a measure' allow their component nouns to be relativized, and others like *te o ireru* 'revise' do not.

(17) a. *[t$_i$ utu]*   *te$_i$*   *ga*   *na-i.*
              strike   hand   NOM   NEG PRS
          'There is no measure to take.'

   b. *#[t$_i$ ireru]*   *te$_i$*   *ga*   *na-i.*
              put.in   hand   NOM   NEG PRS
          'There is no hand to put in.'

Similarly, idioms show a difference in the applicability of nominal modification to their component nouns, as shown in (18).

(18)　a.　*(zyuubun-na)　ki　　o　　kubaru*
　　　　　enough　　　　mind　ACC　deliver
　　　　　'pay (enough) attention to'

　　　b.　*(*zyuubun-na)　ki　　o　　tukeru*
　　　　　enough　　　　　mind　ACC　attach(tr.)
　　　　　'be mindful (enough)'

Note that in (18), if the nominal modifier *zyuubun-na* 'enough' is replaced by an adverbial modifier *zyuubun-ni* 'sufficiently', both idioms are acceptable because it modifies the entire verbal phrases.

　　　Whether adverbial modifiers can be positioned inside idiom chunks depends on their degree of syntactic cohesion. Some idioms allow adverbs to intervene, and others do not.

(19)　a.　*#te　　o　　tadatini　　　tukeru*
　　　　　hand　ACC　immediately　attach(tr.)
　　　　　'start working immediately'

　　　b.　*te　　o　　tadatini　　　utu*
　　　　　hand　ACC　immediately　strike
　　　　　'take a measure immediately.'

(19a) shows that *te o tukeru* 'start (working)' is syntactically frozen and cannot be broken up by an intervening adverb on its idiom reading, whereas (19b) shows that the bond between the noun and verb in *te o utu* 'take a measure' is not so strong, as also shown by relativization (17a). Note that both examples in (19) are equally acceptable if the adverb is placed in front of the idiom chunks.

　　　While applicability of insertion and reordering operations varies according to the degree of cohesion in idiomatic sequences, idioms more often than not retain their syntactic visibility, even when their sequences look fairly tight at first blush. For illustration, let us consider *kote o kazasu* [forearm ACC hold] 'shade one's eye with one hand (to look distantly or to protect the eye from sunlight) [lit. 'hold one's hand up above one's eye']', which Asuka (1982) lists as the tightest idiom. This idiom resists all the syntactic operations he discusses, including passivization and relativization, as (20) shows.

(20)　a.　**Kote　　ga　　kazas-are-ta.*　　　b.　**[t_i　kazasi-ta]　kote_i*
　　　　　forearm　NOM　hold-PASS PST　　　　　　　　hold-PST　　forearm
　　　　　'His hand was held over his eye.'　　　'the hand which he held over his eye'

The ungrammaticality of (20a, b) strongly suggests that the object *kote* (lit.) 'small arm, forearm' is non-referential. Nevertheless, this same idiom allows nominalization, where the case on the noun is changed from accusative to genitive, as in (21).

(21) a. *kote   o   kazasu*   b. *kote   no   kazasi-kata*
　　　forearm   ACC   hold　　　　　forearm   GEN   hold-way
　　　'shade an eye with one hand'　　'the way of shading an eye with one hand'

The appearance of the genitive case in (21b) conforms to the general pattern of *-kata* nominalization, where the arguments of ordinary verb phrases are marked with genitive case, and not with verbal case, as illustrated by (22).

(22) a. *hon   o   yomu*   b. *hon   {no/*o}   yomi-kata*
　　　book   ACC   read　　　　　book   {GEN/*ACC}   read-way
　　　'read books'　　　　　　　'the way of reading books.'

In (22b), the object of *yomu* 'read' receives genitive marking under *-kata* nominalization, and this case marking pattern is shared by the nominalization of the idiom *kote o kazasu* in (21b). This observation shows that the internal structure of the idiom *kote o kazasu* is syntactically visible.

　　A still more remarkable fact is that there are idioms that can (optionally) maintain the original verbal case marker even after they are nominalized. A case in point is the idiom *ki o tukeru* 'pay attention', which allows either accusative or genitive marking on the noun under nominalization.

(23)　a.　*Gakusei   ga   ki   o   tuke-ru.*
　　　　　student   NOM   mind   ACC   attach(tr.)-PRS
　　　　　'The students will pay attention.'

　　　b.　*(gakusei   no)   ki   {o/no}   tuke-kata*
　　　　　student   GEN   mind   {ACC/GEN}   attach(tr.)-way
　　　　　'(the students') way of paying attention'

The fact that the noun *ki* 'attention' can be marked with either accusative or genitive case in (23b) suggests that the entire sequence may serve as either a complex or a single-word predicate. Under the analysis of Kishimoto (2009a), a nominalized verb phrase that retains the original verbal case marking under *-kata* nominalization is considered to have undergone a kind of incorporation to form a single-word predicate. In fact, the appearance of the accusative case in (23b) can be attributed to the fact that the noun is incorporated to the predicate.

An intriguing fact suggestive of the special status of *ki o tukeru* 'pay attention' is found in subject honorification. Observe (24), where the honorific prefix *o-* appears on the noun *ki* 'attention'.

(24) *Itoo-sensei ga   sono koto   ni   **o**-ki   o   tuke-**ni-nat**-ta.*
     Itō-teacher NOM that matter LOC HON-mind ACC attach-DAT become PST
     'Professor Itō paid much attention to that matter.'

(Needless to say, it is also possible to have the honorific prefix on the verb *tukeru* 'attach', as in *ki o **o**-tuke-**ni-naru*** [mind ACC HON-attach(tr.)-DAT-become].) The appearance of the honorific prefix in (24) is irregular in the light of the general rule that the prefix *o-* can only be attached to a verbal element, as seen from the degraded acceptability of (25a) involving the idiom *ki o tukau* [mind ACC use] 'be mindful'.

(25) a. ?**Itoo sensei   ga   sono koto   ni   **o**-ki   o   tukai-**ni-nat**-ta.*
        Itō   teacher NOM that matter LOC HON-mind ACC use-DAT-become PST
        'Professor Itō paid much attention to that matter.'

    b. *Itoo sensei   ga   sono seito   o   **o**-[ki-zukai]-**ni-nat**-ta.*
       Itō   teacher NOM that student ACC HON-[mind-use]-DAT-become PST
       'Professor Itō cared about the student.'

Incidentally, (25b) shows that the subject-honorific prefix can appear in front of the noun *ki* 'mind' if the noun + verb combination is turned into a compound verb. Now, a comparison of (24) with (25a, b) suggests that the apparently irregular honorification in (24) is comparable to the compound verb in (25b) rather than with the verb phrase in (25a). On Kishimoto's (2009a) analysis, the idiom chunk in (24) is reanalyzed as a single verb while maintaining the accusative marking, and the honorific prefix is attached to the whole of the lexicalized verb, as in ***o**-[ki o tuke]*$_\text{V}$-***ni-naru***.

4 Unaccusative and passive idioms

One notable feature of Japanese idioms, which is not observed in English, is that transitive-intransitive pairs of idioms are often available, as in (26).

(26)   a.   *tyatya o ireru* [tea ACC pour] 'interrupt'
            tyatya ga hairu [tea NOM enter] 'be interrupted'

      b.   *ohati o mawasu* [bowl ACC turn(tr.)] 'give a turn'
            ohati ga mawaru [bowl NOM turn(intr.)] 'get a turn'

    c.   *koe o kakeru* [voice ACC hang(tr.)] 'invite'
        koe ga kakaru [voice NOM hang(intr.)] 'get an invitation'

    d.   *tuti o tukeru* [soil ACC attach(tr.)] 'beat'
        tuti ga tuku [soil NOM attach(intr.)] 'lose'

The idioms in (26) retain their idiomatic meanings regardless of the verb's transitivity. Note that the transitive/intransitive pairs like *tukeru* 'attach(tr.)'/*tuku* 'attach (intr.)' and *hairu* 'enter'/*ireru* 'put in' are related by causative alternation, and that in the intransitive idioms, the nominative arguments are included in the idiom sequences. When idioms alternate between transitive and intransitive forms, a concomitant change in meaning takes place, as indicated in (27).

(27)   <u>*tyatya ga hairu*</u>   'X is interrupted'

           ↕

    <u>*tyatya o ireru*</u>   'X causes Y to be interrupted (='X interrupts Y')'

This shows that Japanese idioms can preserve their idiom interpretations regardless of whether or not the 'causative' component is included as part of their meanings. In other words, causative alternation does not affect the possibility of idiom formation.

    English differs markedly from Japanese in this regard. For example, *break* can appear in both transitive and intransitive idioms. The verb *break* is used intransitively in the idiom *All hell broke loose*, but this idiom does not have a transitive counterpart **John broke all hell loose*. Conversely, the verb *break* appearing in the idiom *break the ice* is used only transitively, as in *John broke the ice*, and it is not possible to have an intransitive variant like *The ice broke, although the idiom can be passivized, as in *The ice was broken*. (This idiom does have a corresponding noun, *ice-breaker*, which derives from the transitive idiom.)

    Both transitive and intransitive verbs in (26) are usable as idioms, but this does not mean that causative alternation may take place without restriction. In effect, the availability of idiomatic meanings often varies according to the transitivity of predicates. The idioms shown in (28) are feasible only in the intransitive forms. (The examples marked "#" are acceptable only on the literal meanings.)

(28)   a.   *kuti ga tatu* [mouth NOM stand(intr.)] 'speak fluently'
         #kuti o tateru [mouth ACC stand(tr.)]

       b.   *mizu ga hairu* [water NOM enter] 'suspend'
         #mizu o ireru [water ACC put.in]

       c.   *kokoro ga sizumu* [heart NOM sink(intr.)] 'depressed'
         #kokoro o sizumeru [heart ACC sink(tr.)]

There are also idioms that lack intransitive variants. The idioms listed in (29) can be used only in transitive forms.

(29) a. *otya o nigosu* [tea ACC puddle] 'give an evasive answer'
　　　#otya ga nigoru [tea NOM get.muddy]

　　b. *kuti o hasamu* [mouth ACC enclose] 'intervene'
　　　#kuti ga hasamaru [mouth NOM get.inside]

　　c. *inoti o otosu* [life ACC drop(tr.)] 'die'
　　　#inoti ga otiru [life NOM drop(intr.)]

The important point here is that Japanese has quite a few idioms that alternate between intransitive and transitive forms, a phenomenon that is not observed with English idioms.

　　The availability of idioms is affected by argument alternation; for instance, *nuru* 'paint' is a locative alternation verb, as seen in (30).

(30) a. *Ken　ga　kabe　ni　akai　penki　o　nut-ta.*
　　　Ken　NOM　wall　LOC　red　paint　ACC　paint-PST
　　　'Ken painted red paint on the wall.' [motion frame]

　　b. *Ken　ga　kabe　o　akai　penki　de　nut-ta.*
　　　Ken　NOM　wall　ACC　red　paint　INST　paint-PST
　　　'Ken painted the wall with red paint.' [change-of-state frame]

The idiom *kao ni doro o nuru* 'bring a shame, disgrace [lit. smear mud on someone's face]' can appear in the (30a)-type construction, but not in the (30b)-type construction.

(31) a. *Ken　ga　[sensei　no　kao]　ni　doro　o　nut-ta.*
　　　Ken　NOM　teacher　GEN　face　LOC　mud　ACC　paint-PST
　　　'Ken brought a shame on his teacher.'

　　b. *#Ken　ga　[sensei　no　kao]　o　doro　de　nut-ta.*
　　　Ken　NOM　teacher　GEN　face　ACC　mud　INST　paint-PST
　　　OK only on the literal meaning 'Ken smeared his teacher's face with mud.'

The difference in the availability of the idiomatic meaning comes from the distinct meanings of the two variant constructions. When used non-idiomatically, the verb *nuru* 'paint' can be associated with a 'change of state' meaning or a 'motion' meaning, as shown in (32a) (Kishimoto 2001: see also Okutsu 1981). When used in the idiom *kao ni doro o nuru* 'bring a shame', however, the same verb can fit only into a

motion frame, as shown in (32b). This is because the idiom has a metaphorical motion sense 'X brings a shame to Y'.

(32) a. *X ga Z ni Y o nuru*   'X causes Y to move onto Z'
   X ga Y o Z de nuru   'X causes Z to become painted with Y'

b. *X ga Y no kao ni doro o nuru*
   'X brings a shame to Y' < 'X causes Y to move onto Z'

The idiom in question is not used as a causative verb of change of state meaning 'X causes Y to change a state', because the causer is not taken to exert a direct influence over Y or to take any action directed toward Y. Hence, the idiom can appear in the motion frame in (31a), but not the change-of-state frame in (31b).

One often-discussed constraint imposed on idiom formation is the so-called "adjacency condition", which requires relevant words to form a continuous sequence (at some point of derivation) in order to obtain idiomatic meanings. Japanese and English differ in how the adjacency condition applies. This can be readily confirmed by passivizable idioms. English has a large number of idioms that may be passivized (e.g. *break the ice*, *draw a line*, etc.) (Culicover 1976; Jackendoff 1997; and others).

(33) a. *John broke the ice.*

b. *The ice was broken.*

In English passive idioms, the passive subjects – which constitute part of idiom sequences – are moved to the subject position, and the idiomatic meaning is preserved even after the nominal part of the idiom is separated from the verb as the subject.

(34) [$_{TP}$ The ice [$_{VP}$ was broken ~~the ice~~]]

This fact suggests that in English, the adjacency condition does not have to hold at the surface level, provided that it is satisfied at the underlying level.

Japanese also has transitive idioms that can undergo direct passivization. In fact, Japanese abounds with idioms to which direct passivization may apply while retaining their idiomatic meanings. Some such idioms are listed in (35).

(35) a. *kuti o tukeru* [mouth ACC attach(tr.)] 'have a bite of'

b. *geta o azukeru* [clog ACC leave(tr.)] 'leave a final decision to'

c. *koe o kakeru* [voice ACC hang(tr.)] 'invite'

d. *te o kuwaeru* [hand ACC add] 'add work to, revise'

Example (36b) shows that the argument *mizu* 'water' of the transitive idiom *mizu o sasu* 'interrupt' can be promoted to a passive subject via passivization.

(36) a. *Ken   ga   giron   ni   mizu   o   sasi-ta.*
      Ken   NOM   discussion   LOC   water   ACC   pour PST
      'Ken interrupted the discussion.'

   b. *(Ken   ni-yotte)   giron   ni   mizu   ga   sas-are-ta.*
      Ken   by   discussion   LOC   water   NOM   pour PASS PST
      'The discussion was interrupted (by Ken).'

   c. **Mizu   ga   giron   ni   sas-are-ta.*
      water   NOM   discussion   LOC   pour PASS PST
      'The discussion was interrupted.'

In (36b), the argument *mizu* 'water' that forms part of the idiom *mizu o sasu* is assigned nominative case via passivization. Nevertheless, the nominative argument cannot be moved across the locative argument, i.e. it cannot be moved out of VP, as shown in (36c). In a nutshell, the idiomatic meanings obtain only in the configurations shown in (37).

(37) a. [$_{TP}$ [$_{VP}$ X *ni   mizu   o   sasu*]]
   b. [$_{TP}$ [$_{VP}$ X *ni   mizu   ga   sas-are-ru*]]

In (37b), the passive subject *mizu* (as part of the idiom sequence) must occur in the same position that it occupies in the active sentence with the transitive idiom (37a). This shows that because of the syntactic fixedness of idiom constituents, overt movement of passive subjects is not instantiated even under direct passivization when they constitute part of idioms, i.e. the underlying order of idiom sequence is reflected in the surface form. This fact suggests that in Japanese, the adjacency condition holds at the surface as well as the underlying level.

Not all transitive idioms are passivizable. Just as English has unpassivizable idioms (e.g. *kick the bucket* cannot be passivized, as in **The bucket was kicked*), so Japanese too has unpassivizable transitive idioms, as in (38).

(38) a. *Ken   ga   mitikusa   o   kut-ta.*
      Ken   NOM   weed.on.road   ACC   eat PST
      'Ken idled away.'

   b. *#Ken   ni-yotte   mitikusa   ga   kuw-are-ta.*
      Ken   by   weed.on.road   NOM   eat-PASS PST
      OK only on the literal meaning 'The weeds on the road were eaten by Ken.'

The idiom *mitikusa o ku(w)u* 'idle away' cannot be passivized on the relevant idiomatic meaning. A sample list of unpassivizable transitive idioms is given in (39).

(39) a. *abura o uru* [oil ACC sell] 'idle away one's time'

    b. *taka o kukuru* [amount ACC bind] 'think too lightly'

    c. *te o yaku* [hand ACC burn(tr.)] 'difficult to control'

In Japanese, the adjacency condition needed for idiom formation holds not merely at the underlying level but also at the surface level. This can also be illustrated by the fact that the idiom *mi ni tukeru* 'master, acquire (a skill)', which literally means 'attach to one's body', can have its idiom meaning only when the two elements *mi-ni* and *tukeru* are aligned in this order.

(40) a. *{tokei/eigo}*   *o*   *mi*   *ni*   *tukeru*
       {watch/English} ACC body LOC attach(tr.)
       'wear watch/master English'

    b. *mi*   *ni*   *{tokei/*eigo}*   *o*   *tukeru*
       body LOC {watch/English} ACC attach(tr.)
       'wear a watch/*master English'

In Japanese, nominal constituents can be scrambled fairly freely, often with no semantic effects. Thus, if the sequence *mi ni tukeru* is interpreted non-idiomatically, the order of the arguments may be reversed, as is the case with the object 'watch' in (40b). On the idiom interpretation of 'master (English)', by contrast, the locative argument *mi ni* 'on one's body' cannot be scrambled across the accusative object. This fact of scrambling also suggests that the surface forms of idiom sequences need to reflect the underlying order of their constituents.

In intransitive idioms, such as *koe ga kakaru* [voice NOM hang(intr.)] 'get invited' and *tuti ga tuku* [mud NOM attach(intr.)] 'be beaten, lose', the single idiomatic arguments (i.e. the nominative subjects) are included in the idiom sequences. Observe that the idiom *koe ga kakaru* has a transitive variant *koe o kakeru*.

(41) a. *Ken*   *ga*   *Mari*   *ni*   *koe*   *o*   *kake-ta*.
       Ken NOM Mari LOC voice ACC hang(tr.) PST
       'Ken invited Mari.'

    b. *Mari*   *ni*   *koe*   *ga*   *kakat-ta*.
       Mari LOC voice NOM hang(intr.) PST
       'Mari was invited.'

Notably, the nominative-marked *koe* in the intransitive idiom in (41b) occupies the same position as the accusative-marked *koe* in the transitive idiom in (41a). This fact suggests that the nominative subject remains in the base-generated object position (see Kishimoto 2010). Furthermore, in both variants, the argument *koe* 'voice' included in the idiom sequences cannot be moved away from the verb via scrambling, as illustrated in (42).

(42) a. ?**Ken ga [koe o]ᵢ Mari ni tᵢ kake-ta.*
         Ken NOM voice ACC Mari LOC   hang(tr.) PST
         'Ken invited Mari.'

   b. **[Koe ga]ᵢ Mari ni tᵢ kakat-ta.*
         voice NOM Mari LOC   hang(intr.) PST
         'Mari was invited.'

Both the transitive idiom *koe o kakeru* and its intransitive counterpart *koe ga kakaru* take *ni*-marked arguments, but they are not subjects.[6] Thus, these *ni*-marked arguments cannot be the antecedent of subject-oriented reflexive *zibun* 'self' whereas the nominative argument of the transitive *koe o kakeru* can, as (43) shows.

(43) a. *Kenᵢ ga Marⱼ ni zibunᵢ/*ⱼ no heya-de koe o kake-ta.*
         Ken NOM Mari LOC self   GEN room in voice ACC hang PST
         (lit.) 'Ken invited Mari in self's room.'

   b. *Mariᵢ ni zibun*ᵢ no heya de koe ga kakat-ta.*
         Mari LOC self   GEN room in voice NOM hang PST
         (lit.) 'Mari got an invitation in self's room.'

In a similar vein, the *ni*-marked arguments of the idioms *koe o kakeru* and *koe ga kakaru* cannot be targeted for subject honorification, while the nominative argument of *koe o kakeru* can, as in (44).

(44) a. **Mari ga Itoo-sensei ni koe o o-kake-ni-nat-ta.*
         Mari NOM Itō-teacher LOC voice ACC HON-hang(tr.)-DAT-become PST
         'Mari invited Professor Itō.'

   b. *Itoo-sensei ga Mari ni koe o o-kake-ni-nat-ta.*
         Itō-teacher NOM Mari LOC voice ACC HON-hang(tr.)-DAT-become PST
         'Professor Itō invited Mari.'

---

6 The particle *ni* can be either a locative or dative case marker. In the case of the unaccusative idioms under discussion, *ni* is identified as a locative marker, since the verbs take a *ni* marked locative argument when used in non idiomatic senses.

    c. **Itoo-sensei   ni   koe   ga   o-kakari-ni-nat-ta.*
       Itō-teacher   LOC   voice   NOM   HON-hang(intr.)-DAT-become PST
       'Professor Itō was invited.'

This state of affairs is expected, since the verbs *kakeru* 'hang (tr.)' and *kakaru* 'hang (intr.)' are not stative predicates that can sanction dative subjects (cf. Chapter 16 [Kishimoto, this volume]).

    In (41b), the noun *koe* 'voice' is identified as a subject by the nominative marking, but this argument must be placed to the right of the PP, i.e. in object position. This fact suggests that the subjects constituting part of unaccusative idioms are frozen in their original object position.

(45) a. [$_{TP}$ [$_{VP}$ X *ni* *koe* *o*   *kakeru*]]
      b. [$_{TP}$ [$_{VP}$ X *ni* *koe* *ga* *kakaru*]]

In (45b), *koe ga* [voice NOM] appears in the position where the subject argument of an unaccusative predicate, i.e. the internal argument, appears at the underlying level. The surface ordering requirement imposed on Japanese idioms thus brings about the consequence that the base-generated position of an internal argument forming part of idiom constituents is visible in the surface string.

    Some clausal idioms like *gata ga kuru* [jolt NOM come] 'come close to a breakdown' and *me ga iku* [eye NOM go] 'get attracted' are formed on the basis of unaccusative verbs that do not have transitive counterparts. Even with these unaccusative idioms, the nominative subjects included in the idioms must appear to the right of the *ni*-marked PPs, and cannot be scrambled to the sentence-initial position, as shown in (46).

(46) a. *Konpyuutaa   ni   gata   ga   ki-ta.*
         computer   LOC   jolt   NOM   come PST
         'The computer came close to a breakdown.'

      b. **[Gata   ga]$_i$ konpyuutaa   ni   t$_i$ ki-ta.*
         jolt   NOM   computer   LOC   come PST
         'The computer came close to a breakdown.'

Inasmuch as verbs carrying non-intentional meanings are categorized as unaccusative, the subjects of these idioms appear in object position. Japanese has a large number of unaccusative idioms (without transitive counterparts) that place their nominative subjects inside the idiom sequences, as exemplified in (47).

(47) a. *mizu ga hairu* [water NOM enter] 'get interrupted'

  b. *gata ga kuru* [jolt NOM come] 'come close to a breakdown'

  c. *tuki ga {kuru/mawaru}* [luck NOM {come/turn(intr.)}] 'get lucky'

  d. *ki ga muku* [mind NOM direct] 'get interested'

  e. *hibi ga iku* [crack NOM go] 'get a crack'

All the unaccusative idioms in (47) have internal arguments (marked with nominative case) that cannot be displaced from their original position. Since the underlying order of arguments appearing in idioms is reflected in the surface forms, it comes as no surprise that the nominative argument of *gata ga kuru* 'come close to a breakdown' remains in object position, where an internal argument is generated at the underlying level.

    Note at this point that some idioms involving both locative (*ni*) and accusative arguments have variable ordering, with different meanings. Idioms like *te o ireru* 'revise' in (48a) instantiate the "*ni-o*" case alignment, whereas idioms like *te ni ireru* 'acquire' in (48b) have the reverse "*o-ni*" case alignment.

(48) a. *Ken   wa   genkoo   ni   te   o   ire-ta.*
      Ken   TOP   draft   LOC   hand   ACC   put.in-PST
      'Ken revised the draft.'

  b. *Ken   wa   kuruma   o   te   ni   ire-ta.*
      Ken   TOP   car   ACC   hand   LOC   put.in-PST
      'Ken acquired a car.'

These transitive idioms have intransitive variants, where their two internal arguments are aligned in the same order as those of the transitive counterparts, as seen in (49).[7]

---

7 Of course, there are idioms that do not have intransitive counterparts, as exemplified in (i) and (ii).

(i) a. *Ken   ga   konran   ni   wa   o   kake ta.*
      Ken   NOM   confusion   LOC   ring   ACC   hang(tr.)   PST
      'Ken further complicated confusion.'

  b. **Konran   ni   wa   ga   kakat ta.*
      confusion   LOC   ring   NOM   hang(intr.)   PST
      'Confusion was further complicated.'

(ii) a. *Ken   wa   sainoo   o   hana   ni   kake ta.*
      Ken   TOP   talent   ACC   nose   LOC   hang(tr.)   PST
      'Ken boasted of his talent.'

  b. **Sainoo   ga   hana   ni   kakat ta.*
      talent   NOM   nose   LOC   hang(intr.)   PST
      'Talent was boasted of.'

Some idioms that do not have intransitive variants include accusative arguments in the idiom sequences, as in (i). There are also idioms that include *ni* marked arguments in the idiom sequences, as shown in (ii).

(49) a. *Genkoo   ni   te   ga   hait-ta.*
       draft   LOC   hand   NOM   enter-PST
       'The draft was revised.'

    b. *Kuruma   ga   te   ni   hait-ta.*
       car   NOM   hand   LOC   enter-PST
       'A car was acquired.'

The syntactic behavior of the idioms *te o ireru* and *te ni hairu* suggests that the accusative arguments should be identified as internal arguments base-generated in direct object position.

In (48), the *ni*-marked argument may either precede or follow the accusative argument depending on the type of idiom. On the basis of this fact, Miyagawa and Tsujioka (2004) suggest that the two internal arguments of three-place verbs are freely ordered. Nevertheless, a closer look at the idioms reveals that the syntactic status of *ni*-marked arguments differs depending on the positions where they appear. This can be easily seen by considering *-kata* nominalization.

(50) a. *genkoo   {e   no/*ni}   te   no   ire-kata*
       draft   {DIR   GEN/LOC}   hand   GEN   put.in-way
       'the way of revising the draft'

    b. *kuruma   no   te   {*e   no/ni}   ire-kata*
       car   GEN   hand   {DIR   GEN/LOC}   put.in-way
       'the way of acquiring the car'

When the idiom is nominalized by the suffix *-kata* 'way', the non-idiomatic locative argument ('draft') preceding the body-part noun *te* 'hand' is marked with the genitive case *no* (preceded by *e* 'to', an additional particle to indicate 'direction'), and not with the locative *ni*, as in (50a). On the other hand, the idiomatic *ni*-marked argument in (50b) retains verbal *ni* marking even if the clause is nominalized, and the nominal *e no* marking is not allowed.[8] This difference of case marking in

---

8 The same case marking patterns are observed for the pair of the intransitive counteparts of the idioms, as shown in (i).

(i) a. *genkoo   {e   no/*ni}   te   no   hairi kata*
       draft   {DIR   GEN/LOC}   hand   GEN   enter way
       'the way of revising the draft'

    b. *kuruma   no   te   {*e   no/ni}   hairi kata*
       car   GEN   hand   {DIR   GEN/LOC}   enter way
       'the way of getting the car'

This fact also leads to the conclusion that the two types of *ni* marked arguments have distinct syntactic status.

nominalized clauses would not be expected under Miyagawa and Tsujioka's (2004) assumption that the positions of two internal arguments are free. Rather, it is feasible that the locative argument following an object occupies a special position that can be filled only by an idiom argument, as discussed by Kishimoto (2008). This view gains further support from (51), where the two types of *ni*-marked arguments co-exist.

(51) a. *Ken ga zyoosi ni human o kuti ni dasi-ta.*
  Ken NOM boss LOC complaint ACC mouth LOC let.out-PST
  'Ken made a complaint to his boss.'

  b. *zyoosi {e no/*ni} human no kuti ni dasi-kata*
  boss {DIR GEN/LOC} complaint GEN mouth LOC let.out-way
  'the way of making a complaint to the boss'

  c. *zyoosi e no human no kuti {*e no/ni} dasi-kata*
  boss DIR GEN complaint GEN mouth {DIR GEN/LOC} let.out-way
  'the way of making a complaint to the boss'

The facts regarding *kuti ni dasu* 'say' suggest that there should be two distinct slots that can be filled by independent *ni*-marked arguments, as in [$_{VP}$ X *ni* Y *o* (Z *ni*) V]. Note that the locative slot Z following an accusative argument is reserved for idiomatic arguments. This is confirmed by the fact that while idiomatic arguments in dative-V idioms retain their verbal case marking, non-idiomatic arguments cannot receive verbal case marking under *-kata* nominalization.

In this section, it has been shown that arguments constituting parts of idioms must stay in the position where they are base-generated. The facts of passive and unaccusative idioms show that in Japanese, unlike in English, the adjacency requirement for idiom formation holds at both underlying and surface levels.

5 Possessor-raising idioms

While idioms have fixed sequences of lexical items, they often have free (variable) positions within them, which can be filled by non-idiomatic, referential expressions. Such a variable position is most typically available inside an inalienable body-part nominal, as in (52a).

(52) a. *Kono kaisya ni wa [sensei no kao] ga kik-u.*
  this firm LOC TOP teacher GEN face NOM take.effect-PRS
  'The teacher has strong influence over this firm.'

  b. *Sensei ga kono kaisya ni kao ga kik-u.*
  teacher NOM this firm LOC face NOM take.effect-PRS
  'The teacher has strong influence over this firm.'

In (52a), the inalienable nominal *kao* 'face' contains in it a variable position that does not contribute to the idiom meaning, and this position can be filled by an ordinary noun phrase that represents the "possessor" of *kao*, in a way similar to the variable *X* in English idioms like *The cat's got X's tongue* and *pull X's leg*. Japanese differs from English, however, in that the possessor argument does not have to appear as the modifier of the body-part noun inside the idiom sequence, as in (52b). The synonymy of (52a) and (52b) strongly suggests that the nominative subject argument *sensei ga* [teacher NOM] in (52b) has been moved out from inside the inalienable nominal *kao* by a syntactic process of possessor raising (or possessor ascension), as depicted in (53).

(53)   [X   [X̶   *kao*] *ga kiku*]

Since possessor raising is a syntactic operation that does not affect the meaning of the idiom *kao ga kiku*, (52a) and (52b) share the same interpretation 'have strong influence over', as (54) illustrates.

(54) a.   Y *ni* [X *no kao*] *ga* *kiku* ⟷   'X has strong influence over Y'

  b. X *ga* Y *ni*   *kao* *ga* *kiku* ⟵

In Japanese, possessor raising is attested in a number of different constructions. Observe the synonymy of the pair in (55).

(55) a. *[Ken   no   kuti]   ga   koe-te   i-ru.*
        Ken   GEN   mouth   NOM   fatten-GER   be-PRS
        'Ken is dainty about his food.'

  b. *Ken   ga   kanari   kuti   ga   koe-te   i-ru.*
        Ken   NOM   fairly   mouth   NOM   fatten-GER   be-PRS
        'Ken is dainty about his food.'

(55a) represents the expected genitive construction, whereas (55b) represents the so-called "major subject" construction, where the leftmost nominative phrase (i.e. the major subject) is licensed by a possessor-possessum relation with the second "small" subject denoting the body-part *kuti* 'mouth'. In (55b), the possessor 'Ken' is separate from the possessum 'mouth', as evidenced by the fact that the adverb *kanari* 'fairly' can intervene between the two. These two constructions share the idiomatic meaning regardless of whether the possessor occurs inside or outside the idiom chunk. Similar examples shown in (56) also allow the two variants of the genitive possessor construction and the major-subject construction.

(56) a. *X {ga/no} kuti ga omoi* [X {NOM/GEN} mouth NOM heavy] 'reluctant to talk'

     b. *X {ga/no} kuti ga katai* [X {NOM/GEN} mouth NOM firm] 'not disclose secrets'

     c. *X {ga/no} kuti ga umai* [X {NOM/GEN} mouth NOM good] 'have a smooth tongue'

     d. *X {ga/no} kao ga hiroi* [X {NOM/GEN} face NOM wide] 'know a lot of people'

The rule of possessor raising proposed above is essentially (though not strictly) equivalent to what Kuno (1973) called "subjectivization", which was intended to derive multiple subjects from genitive constructions.

     There are cases where possessor raising fails to apply. In (57), for example, the possessor *oya* 'parent' can bear only genitive case, which shows that possessor raising is not applicable.

(57)   *Ken   wa   oya   {no/*ni/*ga/*o}   sune   o   kazit-te   i-ru.*
         Ken   TOP   parent   {GEN/DAT/NOM/ACC}   knee   ACC   bite-GER   be-PRS
         'Ken is financially dependent on his parents.'

Idioms of this type include *X no kata o motu* [X GEN shoulder ACC hold] 'take sides with', *X no o-kabu o ubau* [X GEN HON-share ACC rob] 'outdo someone in specialty' and *X no me o utagau* [X GEN eye ACC doubt] 'cannot believe', none of which allows possessor raising out of direct objects even in the presence of a variable position within the possessum nominals.[9]

     Although possessor raising introduces a new argument, the added argument may or may not function as a (major) subject (Kishimoto 2009a). Note at this point that in a fairly large number of idioms, the possessor is marked with dative case when it is extracted by possessor raising. These idioms are divided into two classes. In one class of idioms, the raised argument marked with dative case counts as a syntactic subject. A case in point is the dative phrase *sensei ni* [teacher DAT] in the example of (58b), which originates from the possessor position of the body-part noun *kuti* 'mouth', as shown in (58a).

---

9 In a few idioms like *hinsyuku o kau* 'be frowned at', it looks as if possessor raising out of direct objects is possible.

(i)   *Ken   ga   sensei   {no/ni/kara}   hinsyku   o   kat ta.*
      Ken   NOM   teacher   {GEN/DAT/ABL}   disgust   ACC   buy PST
      'Ken is frowned at by the teacher.'

Note, however, that an ablative argument is generally not produced by possessor raising, and that the verb *kau* 'buy' can be a three place predicate taking an ablative/dative argument. Furthermore, a certain idiosyncratic condition seems to apply for the derivation of this type of idiom, because, as far as I know, there is no idiom displaying the pattern of "X NOM Y {NOM/GEN} Z ACC V", which can be taken as an unequivocal case showing that an extra argument is created by possessor raising. Thus, it remains to be seen whether it is appropriate to analyze the dative/ablative argument in (i) as an argument derived by possessor raising.

(58) a. *Sono   ryoori   ga   [sensei   no   kuti   ni]   awa-nakat-ta.*
     that   dish   NOM   teacher   GEN   mouth   LOC   fit-NEG PST
     'The teacher did not like that dish.'

   b. *Sensei   ni   sono ryoori   ga   kuti   ni   awa-nakat-ta.*
      teacher   DAT   that   dish   NOM   mouth   LOC   fit-NEG PST
      'The teacher did not like that dish.'

In another class of idioms, exemplified in (59), the argument extracted by possessor raising is marked with dative case but does not serve as a subject.

(59) a. *Mari   ga   [Ken   no   sewa   ni]   nat-te   i-ru.*
     Mari   NOM   Ken   GEN   care   DAT   become-GER   be PST
     'Mari has been taken care of by Ken.'

   b. *Mari   ga   Ken   ni   sewa   ni   nat-te   i-ru.*
      Mari   NOM   Ken   DAT   care   DAT   become-GER   be-PST
      'Mari has been taken care of Ken.'

Superficially, the idiom *sewa ni naru* behaves in the same way as *kuti ni au*, in that the possessor may be marked with either dative or genitive case. The two classes of idioms share the property that the possessor can appear outside the possessum nominal, but differ as to what argument can be targeted for subject honorification.

(60) a. <u>*Itoo-sensei*</u> <u>*ni*</u>   *sono ryoori   ga   kuti   ni   at-te   irassyara-na-i.*
     Itō-teacher   DAT   that   dish   NOM   mouth   LOC   fit-GER   be.HON   NEG PRS
     'Professor Itō did not like that dish.'

   b. <u>*Itoo-sensei*</u> <u>*ga*</u>   *ano hito   ni   sewa ni   nat-te   irassyar-u.*
      Itō-teacher   NOM   that person   DAT   care   DAT   become-GER   be.HON PRS
      'Professor Itō has been taken care of by that person.'

   c. **Ken   ga*   <u>*Itoo-sensei*</u> <u>*ni*</u>   *sewa ni   nat-te   irassyar-u.*
      Ken   NOM   Itō-teacher   DAT   care   DAT   become-GER   be.HON PRS
      'Ken has been taken care of by Professor Itō.'

With the idiom *kuti ni au* in (60a), the dative possessor extracted by possessor raising is targeted by subject honorification. In contrast, with the idiom *sewa ni naru*, the non-raised nominative argument rather than the dative argument serves as a subject-honorific target, as seen in (60b, c). This shows that in the two classes of idioms, the raised possessor arguments assume different grammatical functions. Japanese has a fairly large number of idioms where extracted arguments act as subjects, some of which are listed in (61).

(61) a. *kuti ni au* [mouth LOC fit] 'favor'

b. *(o-)ki ni mesa-nai* [(HON-)mind LOC call-NEG] 'not like'

c. *ki ni sawaru* [mind LOC harm] 'hurt one's feeling'

d. *ki ni kuwa-nai* [mind LOC eat-NEG] 'not like',

e. *te ni amaru* [hand LOC exceed] 'not controllable'

On the other hand, there are not so many idioms where their extracted arguments do not acquire the status of subjects. Two examples are shown in (62).

(62) a. *ma ni au* [interval LOC meet] 'be in time'

b. *yaku ni tatu* [use LOC stand(intr.)] 'of service to'[10]

The two types of idioms mark their possessor arguments with genitive case when they appear in variable positions internal to the idiom sequences. The possessor arguments are marked with dative case when they are extracted from within the possessum nominal in the idiom sequences via possessor raising. This analysis gains credibility from the behavior of possessor honorification, which is allowed if a genitive possessor is taken to be worthy of respect. To be concrete, consider (63) first.

(63) *{Itoo-sensei/*Ken}*   *no*   ***o*-kao**
       {Itō-teacher/Ken}   GEN   HON-face
       '{Professor Itō's/Ken's} face'

Possessor honorification is legitimate if the genitive possessor refers to an individual worthy of respect (see Harada 1976). This condition holds true for the idioms that

---

10 This idiom has two different uses. When the idiom has an experiential sense, taking a dative experiencer and a nominative theme, it is a transitive stative predicate. In this case, the dative argument serves as a subject. When the idiom has an activity sense, the nominative argument is interpreted as an agent, which serves as a subject. This difference is confirmed by the examples in (i).

(i) a. *Sensei   ni   sono hon   ga   totemo yaku ni   tat te   irassyar u.*
         teacher   DAT   that book   NOM   very   use   LOC   stand(intr.) GER   be.HON-PRS
         'That book was of great use to the teacher.'

b. *Kon kai   wa   sensei   ga   totemo yaku ni   tat te   irassyar u.*
    this time   TOP   teacher   NOM   very   use   LOC   stand(intr.) GER   be.HON-PRS
    'The teacher is making a lot of contributions.'

In (i), the idiom has a stative sense, in which case subject honorification targets the dative experiencer. In (ii), on the other hand, the idiom has an activity meaning, in which case the nominative agent serves as the target of subject honorification. Owing to these two uses, either the dative or nominative argument of the idiom *yaku ni tatu* can be the target for subject honorification.

allow possessor extraction as well. In (64), possessor honorification targeting the non-raised genitive arguments is legitimate.

(64) a. *[Itoo-sensei no   o-kuti]   ga   koe-te   i-ru.*
      Itō-teacher   GEN HON-mouth   NOM fatten-GER   be-PRS
      'Professor Itō is dainty about her food.'

    b. *Sono ryoori ga   [sensei   no   o-kuti]   ni   awa-nakat-ta.*
        that   dish   NOM   teacher GEN HON-mouth   LOC fit-NEG PST
        'The teacher did not like that dish.'

    c. *Mari ga   zutto [sensei no   o-sewa]   ni   nat-te   i-ru.*
        Mari   NOM long   teacher GEN HON-care   DAT become-GER be-PRS
        'Mary has long been taken care of by the teacher.'

The external possessors extracted by possessor raising are also qualified as the targets of possessor honorification, even though they are located in a separate position from the host possessums, as in (65).

(65) a. *Itoo-sensei ga   o-kuti   ga   koe-te   i-ru.*
      Itō-teacher   NOM HON-mouth   NOM fatten-GER   be-PRES
      'Professor Itō is dainty about her food.'

    b. *Sensei   ni   sono ryoori ga   o-kuti   ni   awa-nakat-ta.*
        teacher DAT that dish   NOM   HON-mouth   LOC fit-NEG PST
        'The teacher did not like that dish.'

    c. *Mari ga   zutto sensei   ni   o-sewa   ni   nat-te   ir-u.*
        Mari   NOM long   teacher DAT HON-care   DAT become-GER be PRS
        'Mari has long been taken care of by the teacher.'

The examples in (65) behave in the same way as those in (64) in allowing possessor honorification, a fact that suggests that the external possessors originate from the possessor position inside the possessums (Kishimoto 2013).

    With the idiom *sewa ni naru* 'be taken care of', the dative argument is not a target of subject honorification, as in (60c), but still can be targeted by possessor honorification, as shown in (65c). With the idiom *kuti o tukeru* 'eat', in contrast, subject honorification targeting the nominative subject is legitimate, as shown in (66a), but possessor honorification cannot target the same argument even though the inalienable noun *kuti* 'mouth' is semantically related to it, as shown by the ungrammaticality of (66b).

(66) a. <u>*Itoo-sensei*</u> *ga* *sono ryoori ni* *kuti* *o* *tuke-te* **irassyar-*u*.**
Itō-teacher NOM that dish LOC mouth ACC attach(tr.)-GER be.HON PRS
'Professor Itō is having a bite of that dish.'

b. **<u>Itoo-sensei</u>* *ga* *sono ryoori ni* ***o**-kuti* *o* *tuke-ta.*
Itō-teacher NOM that dish LOC HON-mouth ACC attach(tr.)-PST
'Professor Itō had a bite of that dish.'

Since the verb *tukeru* 'attach(tr.)' in the idiom *kuti o tukeru* take the case frame <NOM, LOC, ACC>, it is capable of licensing all the three arguments (nominative, locative, and accusative) in (66a) without invoking possessor ascension. In other words, no extracted argument appears in the clause. Such being the case, it is plausible to conclude that the subject 'Professor Itō' in (66b) is not a target of possessor honorification (though it can be a target of subject honorification), because it does not originate from within the possessum nominal but is related to the possessum only by a binding relation, as in (67b).

(67) a. [ARG$_i$ [$_{NP}$ ~~ARG$_i$~~ N] PRED]]

b. [ARG$_i$ [$_{NP}$ PRO$_i$ N] PRED]]

A comparison of the data in (65) and (66b) shows that possessor honorification is legitimate when possessor raising is involved, as in (67a), but not when the semantic relation is established by binding, as in (67b).

Let us now turn to another pair of idioms like *ki ga tuku* [mind NOM attach(intr.)] 'notice' and *ki o tukeru* [mind ACC attach(tr.)] 'have an eye on'. There is a considerable difference in meaning between the two variants: the intransitive *ki ga tuku* represents a spontaneous event while the transitive *ki o tukeru* describes the experiencer's intentional commitment. Such a semantic difference is often observed for verbs participating in causative alternation. Compare now the paired examples in (68a) and (68b), where the same number of arguments appears in the clause regardless of the transitivity of the predicate.

(68) a. *Ken* *ga* *sono* *koto* *ni* *ki* *ga* *tui-ta.*
Ken NOM that fact LOC mind NOM attach(intr.)-PST
'Ken noticed that fact.'

b. *Ken* *ga* *sono* *koto* *ni* *ki* *o* *tuke-ta.*
Ken NOM that fact LOC mind ACC attach(tr.)-PST
'Ken had an eye on that fact.'

The idiom *ki ga tuku* in (68a) takes the case frame <NOM, LOC, NOM>, and *ki o tukeru* in (68b) <NOM, LOC, ACC>. In both intransitive and transitive idioms in (68),

the experiencer argument 'Ken' serves as a subject, and can thus be the antecedent of the subject-oriented reflexive *zibun*, as shown in (69).

(69) a. *Ken$_i$ ga zibun$_i$ no kodomo ni ki ga tui-ta.*
    Ken NOM self GEN child LOC mind NOM attach(intr.)-PST
    'Ken noticed his children.'

   b. *Ken$_i$ ga zibun$_i$ no kodomo ni ki o tuke-ta.*
    Ken NOM self GEN child LOC mind ACC attach(tr.)-PST
    'Ken had an eye on his children.'

The intransitive idiom *ki ga tuku* in (68a) takes three arguments, in much the same way that *ki o tukeru* in (68b) does, although their case frames are different. The transitive verb *tukeru* 'attach(tr.)' takes a subject, a direct object, and a locative, so the presence of three arguments in (68b) and (69b) is well expected. In contrast, the intransitive *tuku* 'attach(intr.)', which selects two arguments (nominative and locative) in its literal usage, is associated with three arguments (nominative, locative, and another nominative) in (68a) and (69a). Some idioms that behave in the same way as the pair of the idioms *ki ga tuku* and *ki o tukeru* are given in (70).[11]

(70) a. *hara o tateru* [stomach ACC stand(tr.)] 'get angry'
    hara ga tatu [stomach NOM stand(intr.)] 'be angry'

   b. *ki o mawasu* [mind ACC turn(tr.)] 'care too much about'
    ki ga mawaru [mind NOM turn(intr.)] 'be attentive to (small details)'

   c. *me o mukeru* [eye NOM turn(tr.)] 'get interested in'
    me ga muku [eye NOM turn(intr.)] 'have an interest in'

With the intransitive idioms listed in (70), the idiom constituent contains an idiomatic argument identified as a subject by the nominative marker. This argument remains in the base object position, which indicates that the clause-subject position has not been filled. Thus, the experiencer (possessor) argument can be placed in the subject

---

11 The pair of the idioms *onaka o sukasu* 'get hungry' and *onaka ga suku* 'be hungry' shows a similar pattern, in that an extra experiencer argument is added to the intransitive version.

(i) a. *Kodomo ga onaka o sukasi ta.*
    child NOM stomach ACC reduce(tr.) PST
    'The child got hungry.'

   b. *Kodomo ga onaka ga sui ta.*
    child NOM stomach NOM reduce(intr.) PST
    'The child was hungry.'

Since the verb does not take a locative argument, the intransitive idiom has the case frame <NOM, NOM>, while the transitive idiom has the <NOM, ACC> pattern.

position via possessor raising. The addition of an experiencer argument via possessor raising is not possible with the transitive version of the idioms because they have external arguments that fill the subject position. Since the intransitive variants have one more argument than is lexically required by the verbs, it is plausible to assume that the experiencer arguments have originated as the possessors of inalienable body-part nouns. That is, possessor raising brings about the effect of adding extra clausal arguments, which are semantically related to the inalienable nouns.

Theoretically, the existence of such alternating idioms as *ki ga tuku* and *ki o tukeru* – which take three arguments regardless of the transitivity of the base verb – is expected if an intransitive idiom can have an extra experiencer argument furnished by possessor raising.[12] Nevertheless, the experiencer argument of *ki ga tuku* cannot be placed in the possessor position nor can it license possessor honorification, as shown in (71).

(71) a. **Sono koto ni [Ken no ki] ga tui-ta.*
that fact LOC Ken GEN mind NOM attach(intr.)-PST
'Ken noticed that fact.'

b. **Ken ga sono koto ni **o**-ki ga tui-ta.*
Ken NOM that fact LOC HON-mind NOM attach(intr.)-PST
'Ken noticed that fact.'

The inalienable noun *ki* 'mind' allows possessor honorification in cases where the possessor phrase referring to a person worthy of respect is base-generated in its possessor position, as shown by the idiom *ki ni sawaru* 'not like' in (72).

(72) *Sore ga sensei no **o**-ki ni sawat-ta.*
that NOM teacher GEN HON-mind LOC harm-PST
'The teacher did not like that.'

---

12 It is also worth noting that idioms often do not allow additional experiencers to appear as their clausal arguments, as in (i).

(i) a. *(*Ken ga) Mari ni ohati ga mawat ta.*
Ken NOM Mari LOC bowl NOM turn(intr.) PST
'For Ken, Mari took a turn.'

b. *Ken ga Mari ni ohati o mawasi ta.*
Ken NOM Mari LOC bowl ACC turn(tr.) PST
'Ken gave Mari a turn.'

The idiom *ohati ga mawaru* 'take a turn', which has a transitive counterpart *ohati o mawasu* 'give a turn', does not allow an experiencer argument to appear, perhaps owing to the fact that it does not include a body part noun.

The unacceptability of the examples in (71a) and (71b) suggests then that with the idiom *ki ga tuku*, the experiencer ('Ken') does not come from within the possessum phrase but instead is base-generated in the subject position.

Note that some idioms such as *me ni tomeru* [eye LOC stop(tr.)] 'notice' and *me ni tomaru* [eye LOC stop(intr.)] 'notice' display intermediate properties.

(73) a. *Sensei   ga   sore   o   me   ni   tome-ta.*
        teacher NOM that ACC eye LOC stop(tr.)-PST
        'The teacher noticed that.'

   b. *Sore   ga   sensei   no   me   ni   tomat-ta.*
        that NOM teacher GEN eye LOC stop(intr.)-PST
        'The teacher noticed that.'

   c. **Sensei   {ni/ga}   sore   ga   me   ni   tomat-ta.*
        teacher {DAT/NOM} that NOM eye LOC stop(intr.)-PST
        'The teacher noticed that.'

The transitive idiom *me ni tomeru* in (73a) has an experiencer ('teacher') in subject position, whereas the intransitive idiom *me ni tomaru* in (73b) has the same argument inside the possessum nominal. The intransitive idiom *me ni tomaru* does not allow the possessor argument to be extracted from within the possessum nominal, as shown in (73c). With this pair of the idioms, the following contrast in acceptability is observed with regard to possessor honorification.

(74) a. **Sensei   ga   sore   o   **o**-me   ni   tome-ta.*
        teacher NOM that ACC HON-eye LOC stop(tr.)-PST
        'The teacher noticed that.'

   b. *Sore   ga   sensei   no   **o**-me   ni   tomat-ta.*
        that NOM teacher GEN HON-eye LOC stop(intr.)-PST
        'The teacher noticed that.'

The intransitive idiom allows a possessor to appear in the possessum, while the transitive counterpart does not. Granted that possessor raising makes an extra argument available for intransitive idioms, whose subject position is not filled by an argument of the verb, it is reasonable to generalize that possessor raising also serves as a source for the extra experiencer argument associated with the idiom *ki ga tuku* in (68a), even though the effect of possessor raising is no longer visible in this particular idiom.

6 Negative idioms

Yet another remarkable feature of Japanese idioms, noted by Miyaji (ed.) (1982), Miyaji (1999) and others, is that there is a large inventory of idioms that are usable only in the negative form.

(75) a. *te ni oe-nai* [hand LOC can.carry NEG] 'cannot control'

    b. *ki ni kuwa-nai* [mind LOC eat NEG] 'not like',

    c. *simatu ni oe-nai* [disposal LOC can.carry NEG] 'out of control'

Some idioms contain negative polarity items (NPI), which can occur only in negative contexts, with the result that they cannot appear in the affirmative form. Others do not contain any NPIs, but are nevertheless usable only in the negative forms.

    An interesting fact is that many negative idioms behave syntactically as adjectives rather than as verbs. The idioms in (75) can be embedded under *omou* 'think', which takes an adjectival predicate in the small clause complement, suggesting the negative *nai* included in the idioms has the lexical property as an adjective.

(76) a. *Watasi wa [kono kodomo o te ni oe-naku] omo-u.*
        1.sg TOP this child ACC hand LOC can.carry-NEG think-PRS
        'I think this child to be out of control.'

    b. *Watasi wa [kono kodomo o ki ni kuwa-naku] omo-u.*
        1.sg TOP this child ACC mind LOC eat-NEG think-PRS
        'I think this child not to be my favorite.'

    c. *Watasi wa [kono kodomo o simatu ni oe-naku] omo-u.*
        1.sg TOP this child ACC disposal LOC can.carry-NEG think-PRS
        'I think this child to be out of control.'

There is in fact a sense in which the idioms in (75) should behave as adjectives, if lexical adjective *nai* takes the verb as its complement, i.e. they possess the predicate structure [$_{ADJ}$ [$_V$ V] *nai*]. The fact that the idioms in (75) are construed as adjectival idioms is related to the question of how morphologically complex negative adjectives like *tumara-nai* 'boring' are derived (Kishimoto 2009b). This is because the negative adjectives should be derived by way of the reduction of the lexical word *nai* to an affix *nai* via grammaticalization involving the general reduction process: word > clitic > affix (see e.g. Hopper and Traugott 1993). Given this, it is interesting to see whether the three negative idioms in (75) involve different types of negative *nai*.

    Japanese is an agglutinative language where many predicative elements occurring after a main verb function as bound morphemes, forming a tight morphological unit with the verb, regardless of their syntactic status. Owing to this property, differences in the status of *nai* between the three idioms in (75) are not visible in the surface strings. Nevertheless, they can be clearly distinguished by certain syntactic operations. In particular, the distinction of "words", "clitics", and "affixes" assumed in theories of grammaticalization holds neatly for the three instances of *nai* in (75). The idiom *te ni oe-nai* in (75a) represents a case where the negative *nai* functions

syntactically as an independent adjective (i.e. word), despite the fact that it is morphologically dependent on the preceding verb because of the agglutinative morphology. By contrast, the negative *nai* appearing in *ki ni kuwa-nai* in (75b) can be identified as a clitic, which is a dependent element but is visible to the syntax. Finally, negative *nai* in the idiom *simatu ni oe-nai* 'be out of control' in (75c) can be regarded as an affix that derives a single lexical adjective by attaching to the verb *oe(ru)*.

The distinction of the three types of negative *nai* can be shown as follows (see Kishimoto 2009b). First, whether or not *nai* serves as an independent word can be determined according to whether it can be detached from the verb.

(77) a. *Ken   ni   kodomo   ga   te   ni   oe-soo   ni   **na-i.**
       Ken   DAT   child   NOM   hand   LOC   can.carry-likely   COP   NEG PRS
       'Ken is not likely to be able to control the child.'

   b. **Ken   ni   sore   ga   ki   ni   kui-soo   ni   **na-i.**
        Ken   DAT   that   NOM   mind   LOC   eat-likely   COP   NEG PRS
        'Ken is not likely to like that.'

   c. **Ken   wa   simatu   ni   oe-soo   ni   **na-i.**
        Ken   TOP   disposal   LOC   can.carry-likely   COP   NEG PRS
        'Ken is likely to be out of control.'

Among the three idioms in (75), only the idiom *te ni oe-nai* allows the insertion of *soo-da* 'be likely' between the verb and *nai*, showing that the two elements are syntactically separate elements. The instances of *nai* appearing in the idioms *ki ni kuwa-nai* 'not like' and *simatu ni oe-nai* 'out of control' do not qualify as syntactically independent words, because they cannot undergo Neg-separation.

Second, the syntactic visibility of the negative constituents may be measured by looking at whether a simple negative form *nai* can be replaced with another negative form *nai-de iru* 'be not to'. This is because *nai-de iru* is a special negative form that attaches to the verb immediately preceding it.

(78) a. *Ken wa imadani sono kodomo ga   te   ni   oe-**nai-de**   **i-ru.**
       Ken TOP still   that child   NOM   hand   LOC   can.carry-NEG GER   be-PRS
       'Ken is still unable to control that child.'

   b. *Kodomo wa imadani sono koto ga   ki   ni   kuwa-**nai-de** **i-ru.**
        child-TOP   still   that fact NOM   mind   LOC   eat-NEG GER   be-PRS
        'The child still does not like that fact.'

   c.**Ken wa imadani simatu   ni   oe-**nai-de**   **i-ru.**
        Ken TOP still   disposal LOC   can.carry-NEG GER   be PRS
        'Ken has still been out of control.'

The idiom *te ni oe-nai*, which contains a syntactically independent *nai*, is subject to *nai-de iru* replacement, which is also available with the idiom *ki ni kuwa-nai* even though Neg-separation is not allowed for this idiom. This observation indicates that *nai* included in *ki ni kuwa-nai* functions as a clitic, which is morphologically dependent on the preceding verb but is visible syntactically. *Simatu ni oe-nai* exhibits the same behavior as the single-word adjective *tumara-nai* 'boring' in that neither Neg-separation nor *naide iru* replacement is allowed, as shown in (79).

(79) a. **Sono   eiga   wa   tumari-soo   ni   **na-i**.*
that   movie   TOP   clog-likely   COP   NEG PRS
'That movie is not likely to be interesting.'

   b. **Sono   bangumi   wa   imadani   tumara-**nai-de**   **i-ru**.*
that   program   TOP   still   clog-NEG GER   be-PRS
'That program has still been boring.'

In Japanese, an ordinary negator is a grammatical marker (functional category) different from the lexical category of adjective. Negative *nai* appearing in idioms sometimes counts as a functional negator, as in *narihuri kamawa-nai* 'jockey for position, behave selfishly without caring about one's appearance'. This idiom cannot be embedded under *omou* 'think', as in (80), showing that the negative *nai* here does not have the status of a lexical word (i.e. lexical adjective).

(80) **Watasi   wa   [Mari   o   narihuri   kamawa-naku]   omo-u.*
1.sg   TOP   Mari   ACC   appearance   care-NEG   think-PRS
'I think Mari is jockeying for position.'

In addition, the idiom *narihuri kamawa-nai* is amenable to both Neg-separation and *nai-de iru* replacement, suggesting that the negator *nai* behaves as a syntactically independent word.

(81) a. *Ken   wa   narihuri   kamai-soo   ni   **na-i**.*
Ken   TOP   appearance   care-likely   COP   NEG PRS
'Ken is not likely to be jockeying for position.'

   b. *Ken   wa   imadani   narihuri   kamawa-**naide**   **i-ru**.*
Ken   TOP   still   appearance   care-NEG   be-PRS
'Ken still has been jockeying for position.'

The data illustrate that the idiom *narihuri kamawa-nai* includes the negative *nai* acting as a syntactically independent functional word rather than as an adjective.

To recapitulate, the negative morphemes appearing in idioms can be classified into three different types: word, clitic, and affix. The negative *nai* in *te ni oe-nai* (75a) is a syntactically independent lexical word; the same negative morpheme in *ki ni kuwa-nai* (75b) behaves as a clitic, which needs to be attached to its host verb; finally, the *nai* in *simatu ni oe-nai* (75c) is identified as a derivational suffix. The negative *nai* appearing in *narihuri kamawa-nai*, in contrast, is a grammatical negator, which is devoid of adjectival properties.[13]

7 Conclusion and future research perspectives

In this chapter, general and language-specific properties of Japanese idioms have been discussed. Japanese idioms, like those in other languages, have non-compositional meanings, so they must be registered in the lexicon. Nevertheless, idioms often have constituent structure transparent to the syntax, and display syntactic behavior that is not shared with non-idiomatic expressions. Unaccusative, passive, possessor-raising, and negative idioms all possess unique properties, which can easily be detected when they are compared with English idioms. A particularly notable feature of Japanese idioms is that the underlying constituency of idiomatic sequences is reflected in their surface strings, as can be seen by the fact that the internal arguments included in idiom sequences are realized in object position without movement. This property of Japanese idioms allows us to verify the unaccusative hypothesis. Unaccusative idioms give us a good indication that the sole arguments of unaccusative verbs are base-generated in object position, because the nominative arguments of unaccusative idioms, which form part of idiom sequences, appear in the surface object position, without undergoing NP movement.

In the literature on Japanese, a fairly large number of works mention certain peculiar properties of idioms that can be easily discerned when they are compared with ordinary expressions. Nonetheless, very little attention has been paid to the question of what unique properties Japanese idioms have in comparison with idioms in other languages. Even a cursory look at a few cases of Japanese and English idioms reveals that Japanese idioms show a number of differences that distinguish

---

13 The inventory of the three classes of idioms is not small (see Kishimoto 2009b for a list of such idioms). There is good reason to believe that the process of morphologization going from an inde pendent word to the lexical affix involves a diachronic change. The first attested appearance of an idiom containing the lexical word *nai*, i.e. *rati ga aka nai* [fence NOM open-NEG] 'get nowhere', dates back to 1604 1608. This is about the same time that the first use of *nai* as a grammatical marker (combined with verbs) is attested. The clitic type of *nai* is relatively old, and the earliest citation for *ki ni kuwa nai* 'not like' goes back to 1671. Completely lexicalized idioms are fairly recent, and the oldest one, which is *aita kuti ga husagara nai* [open mouth NOM shut-NEG] 'be amazed', dates back to 1700. [The attested dates are based on *Nihon kokugo daijiten* (2000 2002).]

them from English idioms. This leads to a further research question of how and why such differences should arise between the two languages. There are many other aspects of Japanese idioms well worth exploring from a cross-linguistic or comparative perspective. Further extensive research on idioms will likely provide some new insights that enable us to further our understanding of why idioms behave in the way they do.

Acknowledgments

References

Asuka, Hiro'omi. 1982. Nihongo dōshi kan'yōku no kaisōsei [Hierarchy of Japanese verbal idioms]. *Gengo* 11(13), *Tenth Anniversary Special Issue*. 72–81(+100).

Brinton, Laurie and Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.

Culicover, Peter. 1976. *Syntax*. New York: Academic Press.

Everaert, Martin, Erik-Jan ver der Linden, André Schenk and Rob Schreuder (eds.). 1995. *Idioms: Structural and psychological perspectives*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6. 22–42.

Harada, S. I. 1976. Honorifics. In Masayoshi Shibatani (ed.), *Syntax and semantics* 5: *Japanese generative grammar*, 499–561. New York: Academic Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

Ishida, Pricilla. 2000. Dōshi-kan'yōku-nitaisuru tōgosōsa no kaisōsei [On the hierarchy of syntactic operations applicable to verb idioms]. *Nihongokagaku* 7. 24–43.

Ishida, Pricilla. 2004. Dōshi-kan'yōku no imitekikoteisei-o hakaru hōhō [Ways of measuring the semantic fixedness of verbal idioms]. *Kokugogaku* 55. 42–56.

Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Kageyama, Taro. this volume. Chapter 7: Noun-compounding and noun incorporation. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kishimoto, Hideki. 2001. Locative alternation verbs in Japanese: A case study in the interaction between syntax and lexical semantics. *Journal of Japanese Linguistics* 17. 59–81.

Kishimoto, Hideki. 2008. Ditransitive idioms and argument structure. *Journal of East Asian Linguistics* 17. 141–179.

Kishimoto, Hideki. 2009a. Syntactic status of idiomatic subjects in Japanese. *Nanzan Linguistics* 5. 41–60. Nanzan University.

Kishimoto, Hideki. 2009b. On the formation of lexicalized negative adjectives. In Yoko Yumoto and Hideki Kishimoto (eds.), *Goi no imi to bunpō* [Lexical meanings and grammar], 47–64. Tokyo: Kurosio Publishers.

Kishimoto, Hideki. 2010. Subjects and constituent structure in Japanese. *Linguistics* 48. 629–670.

Kishimoto, Hideki. 2013. Covert possessor raising in Japanese. *Natural Language and Linguistic Theory* 31. 161–205.

Kishimoto, Hideki. this volume. Chapter 16: Stative and existential/possessive Predicates formation processes. In Taro Kageyama and Hideki Kishimoto (eds.), *Handbook of Japanese lexicon and word formation*. Berlin: De Gruyter Mouton.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Miyagawa, Shigeru and Takae Tsujioka. 2004. Argument structure and ditransitive verbs in Japanese. *Journal of East Asian Linguistics* 13. 1–38.

Miyaji, Yutaka (ed.). 1982. *Kan'yōku no imi to yōhō* [The usage and meaning of idioms]. Tokyo: Meiji Shoin.

Miyaji, Yutaka. 1999. *Keigo kan'yōku* hyōgen-ron [On honorific and idiomatic expressions]. Tokyo: Meiji Shoin.

Nihon kokugo daijiten [The grand dictionary of the Japanese language] 2000–2002. 2nd edition. Tokyo: Shōgakukan.

Nunberg, Geoffrey, Ivan Sag and Thomas Wasow. 1994. Idioms. *Language* 70. 491–538.

Okutsu, Keiichiro. 1981. Idō-henka-dōshibun [Sentences with motion-change verbs]. *Kokugogaku* 127. 21–32.

Subject index